FIRST-PRINCIPLES THEORY, DISCOVERY, AND DESIGN OF NONLINEAR OPTICAL MATERIALS

A Dissertation

by

HUA WANG

Submitted to the Office of Graduate and Professional Studies of Texas A&M University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee, Xiaofeng Qian
Committee Members, Tahir Çagin
Hong-Cai Joe Zhou
Lei Fang
Head of Department, Ibrahim Karaman

May 2020

Major Subject: Materials Science and Engineering

Copyright 2020 Hua Wang
Nonlinear light-matter interaction plays a key role in the understanding, probing, and ultimately controlling light and matter. In particular, materials with strong nonlinear optical responses are highly desirable for many scientific disciplines and technological applications e.g. ultrafast nonlinear optics, nonlinear biosensing and imaging, efficient generation of entangled photon pairs for quantum computing and quantum sensing, and all-optical transistor and computer. This thesis focuses on the theoretical study and fundamental understanding of nonlinear light-matter interaction using group theory and first-principles electronic structure theory. The developed theoretical framework can also be used for the discovery and design of nonlinear optical materials.

Chapter I of this dissertation discusses the general background and concepts of nonlinear light matter interactions and the motivation of this study. Chapter II provides a general microscopic theory of various nonlinear responses of materials as well as symmetry principles that govern the fundamental nonlinear responses. In addition, density matrix formalism and Floquet theory as well as tight-binding method will be discussed. Chapter III will present the discovery of 2D multiferroic and topological materials.

Chapter IV, V, and VI will discuss several types of nonlinear light-matter interaction in different materials including semiconductors and metals as well as topological materials. Chapter IV focuses on the large second harmonic generation in two-dimensional materials, including first-principles density functional theory approach and the corresponding microscopic mechanisms of second harmonic generation in the low-dimensional materials. Chapter V focuses on the shift photocurrent and circular photocurrent in ferroelectric semiconductors. Such photocurrent is distinct from conventional linear photocurrent as they exhibit nonreciprocal behavior whose current flow direction can be controlled by ferroelectric polarization. Chapter VI focuses on the theory of ferroelectric nonlinear anomalous Hall effect in semimetals and topological materials, and a theoretical prediction of even-odd layer oscillation of ferroelectric nonlinear anomalous Hall effect in few-layer topological semimetals is discussed.
Chapter VII elaborates future directions and important issues to be addressed. In particular, other types of nonlinear light matter interactions are discussed, including surface second harmonic generation, inverse Faraday effect, nonlinear magnetophotocurrent in topological materials etc. In addition to the second-order effects, higher order effects will be also discussed, including third-harmonic generation, two-photon absorption, etc. The results presented demonstrate that the fundamental understanding of the nonlinear optical phenomena together with their symmetry principles may offer unprecedented opportunities for the discovery and design of nonlinear optical materials and enable novel devices such as nonlinear quantum electronics, spintronics, magneto-electronics, and dynamic quantum materials which may foster the second quantum evolution.
DEDICATION

To my mother, my father, my sister, and my grandmother, for their endless support and love.
ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my Ph.D. advisor and committee chair Dr. Xiaofeng Qian for his constant support, crucial guidance, and very valuable mentoring throughout my Ph.D. study. I would like to thank Dr. Tahir Çagin, Dr. Hong-Cai Joe Zhou, and Dr. Lei Fang for their constructive comments and thoughtful suggestions. I am also grateful to Dr. Pokrovsky, a great theoretical physicist, for sharing his rich expertise. Furthermore, I would like to acknowledge the help and great support from the Department of Materials Science and Engineering at Texas A&M University, the financial support from the National Science Foundation, the support and award from 2019 Kaneka Scholarship and 2017 MSEN Best Paper Graduate Student Award, and the advanced computing resources provided by Texas A&M High Performance Research Computing. Finally, I want to take this opportunity to express great gratitude to all my friends for their help and support.
CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Xiaofeng Qian (chair of committee) and Tahir Çagin of the Department of Materials Science and Engineering and Professors Hong-Cai Joe Zhou and Lei Fang of the Department of Chemistry.

The $k \cdot p$ model in Chapter 3 was done by Drs. Chen Fang, Junwei Liu, and Liang Fu.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by the National Science Foundation (NSF) under award number DMR-1753054. Portions of the research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLO</td>
<td>nonlinear optics / nonlinear optical</td>
</tr>
<tr>
<td>DFT</td>
<td>density functional theory</td>
</tr>
<tr>
<td>TDDFT</td>
<td>time-dependent density functional theory</td>
</tr>
<tr>
<td>TB</td>
<td>tight-binding</td>
</tr>
<tr>
<td>WFs</td>
<td>Wannier functions</td>
</tr>
<tr>
<td>2D</td>
<td>two-dimensional</td>
</tr>
<tr>
<td>vdW</td>
<td>van der Waals</td>
</tr>
<tr>
<td>SHG</td>
<td>second-harmonic generation</td>
</tr>
<tr>
<td>SFG</td>
<td>sum-frequency generation</td>
</tr>
<tr>
<td>SPDC</td>
<td>spontaneous parametric down conversion</td>
</tr>
<tr>
<td>SC</td>
<td>shift current</td>
</tr>
<tr>
<td>MSC</td>
<td>magnetic shift current</td>
</tr>
<tr>
<td>CC</td>
<td>circular current</td>
</tr>
<tr>
<td>IC</td>
<td>injection current</td>
</tr>
<tr>
<td>MIC</td>
<td>magnetic injection current</td>
</tr>
<tr>
<td>BPVE</td>
<td>bulk photovoltaic effect</td>
</tr>
<tr>
<td>CPGE</td>
<td>circular photogalvanic effect</td>
</tr>
<tr>
<td>SOC</td>
<td>spin-orbit coupling</td>
</tr>
<tr>
<td>BZ</td>
<td>Brillouin zone</td>
</tr>
<tr>
<td>MOKE</td>
<td>magneto-optic Kerr effect</td>
</tr>
<tr>
<td>MOFE</td>
<td>magneto-optic Faraday effect</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>CONTRIBUTORS AND FUNDING SOURCES</td>
<td>vi</td>
</tr>
<tr>
<td>NOMENCLATURE</td>
<td>vii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xxi</td>
</tr>
</tbody>
</table>

1. INTRODUCTION TO NONLINEAR OPTICS ........................................... 1
   1.1 The Birth of Nonlinear Optics ........................................... 1
   1.2 What Is Nonlinear Optics? ............................................... 1
   1.3 Motivation ........................................................................ 5

2. QUANTUM THEORY OF NONLINEAR OPTICS ......................................... 7
   2.1 A Classical Model of Second-Order Electrical Nonlinearities ........ 7
   2.2 Density Matrix Formalism and Time Dependent Perturbation Theory .... 8
       2.2.1 The Interaction Hamiltonian in Second Quantization ............ 8
       2.2.2 The Polarization and Current Density Operator .................. 12
       2.2.3 Time Dependent Perturbation Theory ................................ 14
   2.3 Floquet Theory, Green’s function and TDDFT ............................ 21
   2.4 Tight-Binding Method and Wannier Representation ......................... 29
   2.5 Group Theory .................................................................. 31

3. DISCOVERY OF 2D MULTIFERROICS AND TOPOLOGICAL MATERIALS .............. 34
   3.1 Two-Dimensional Multiferroics in Monolayer Group IV Monochalcogenides 34
       3.1.1 Brief Introduction to 2D Materials and Multiferroics ............ 34
       3.1.2 Ferroelastic Order and Spontaneous Strain in Monolayer MX ........ 36
       3.1.3 Ferroelectric Order and Spontaneous Polarization in Monolayer MX .... 39
       3.1.4 Ferroelectric Domain Wall in Monolayer MX .......................... 43
3.1.5 Electronic Structure of Ferroelectric Monolayer MX ........................................ 48
3.1.6 Device Concepts Based On 2D Multiferroic Monolayer MX .......................... 50
3.2 Van der Waals Stacking Induced Topological Phase Transition in Layered Ternary
Transition Metal Chalcogenides ........................................................................ 54
3.2.1 Brief Introduction to Topological Materials ....................................................... 54
3.2.2 Prediction of New Class of Topological Materials .............................................. 56
3.2.3 $k \cdot p$ Model .................................................................................................... 67
3.2.4 Experimental Synthesis and Mechanical Exfoliation ..................................... 70

4. OPTICAL SECOND HARMONIC GENERATION ........................................... 75
4.1 Second Harmonic Generation in Two-Dimensional Materials ............................ 75
4.2 First-Principles Calculation of Second Harmonic Generation .............................. 77
4.3 Second Harmonic Generation in 2D Multiferroic Monolayer MX ....................... 81
4.4 Conclusions ........................................................................................................ 92

5. NONLINEAR PHOTOCURRENT RESPONSES .............................................. 93
5.1 Introduction to Nonlinear Photocurrents ............................................................. 93
5.2 Theory of Shift Current and Circular Current ..................................................... 96
5.3 First-Principles Atomistic and Electronic Structure Calculation Methods .......... 99
5.4 First-Principles Calculations and Group Theoretical Analysis of Nonlinear Photo-
current Responses ...................................................................................... 101
5.5 Ferroicity-Driven Nonlinear Photocurrent Switching ........................................ 107
5.6 Generalization of Nonlinear Photocurrent Switching ....................................... 111

6. NONLINEAR ANOMALOUS HALL EFFECT ................................................ 113
6.1 Introduction to Nonlinear Anomalous Hall Effect ............................................. 113
6.2 Second-Order DC Current ................................................................................. 115
6.3 First-Principles Tight-Binding Calculations ...................................................... 119
6.4 Ferroelectric Transition in Bilayer and Trilayer WTe$_2$ .................................... 121
6.5 NAHE in Bilayer and Trilayer WTe$_2$ upon Ferroelectric Switching ................... 125
6.6 Group Theoretical Analysis of NAHE in Bilayer and FNAHE in Trilayer WTe$_2$ .... 130

7. FUTURE DIRECTIONS AND CONCLUSIONS ............................................. 138
7.1 Surface Second Harmonic Generation ............................................................. 138
7.2 Inverse Faraday Effect and Optical Rectification ............................................. 142
7.3 Nonlinear Magnetophotocurrent Responses ..................................................... 144
7.4 Nonlinear Current Response with External Magnetic Field .............................. 150
7.5 Higher Order Nonlinear Optical Effects and Many-Body Effects ..................... 151
7.6 Conclusions ....................................................................................................... 154

REFERENCES ................................................................................................. 155

APPENDIX A. NOTATIONS AND IDENTITIES .............................................. 181
A.1 The Representation of Position Operator .................................................. 181
A.2 Derivation of Commutative Relations ..................................................... 181
A.3 Levi-Civita Symbol and Cross Product .................................................. 183
A.4 Covariant Derivative and Sum Rules ...................................................... 184
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 A Feynman diagram for photon-photon scattering $\gamma_1\gamma_2 \rightarrow \gamma_3\gamma_4$ with electron-positron pair $e^-e^+$</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Typical NLO effects: second harmonic generation, linear electro-optic effect, static polarization (current, current injection), third harmonic generation, quadratic electro-optic effect, and two-photon absorption. $E$ is a static electric field.</td>
<td>4</td>
</tr>
<tr>
<td>2.1 A binary tree diagram of perturbative density operators including intraband (blue node) and interband (orange node) terms. At each node, the density operator is separated into the left child (intraband) and the right child (interband).</td>
<td>17</td>
</tr>
<tr>
<td>3.1 Structure of monolayer Group IV monochalcogenides (MX) and their ferroelastic and ferroelectric orders. (a) Perspective view of free-standing MX monolayer using GeSe as an example. (b) and (c): two ferroelectric states with opposite spontaneous polarization along the $x$ axis: $-P_x$ and $+P_x$. (d) and (e): two ferroelectric states with opposite spontaneous polarization along the $y$ axis: $-P_y$ and $+P_y$. (b,c) and (d,e): two ferroelastic states with spontaneous tensile strain along $x$ and $y$ relative to the centrosymmetric paraelastic structure. The corresponding transformation strain matrices are denoted by $\eta_x$ and $\eta_y$.</td>
<td>36</td>
</tr>
<tr>
<td>3.2 Ferroelastic order and spontaneous strain in monolayer GeSe. (a) Total energy of monolayer GeSe with different lattice parameters $a$ and $b$ with respect to their ferroelectric ground state. Solid green line: phase boundary between paraelectric and ferroelectric phases. Red dashed line: phase boundary between two ferroelastoelectric phases with spontaneous strain along $x$ and $y$, respectively, where ferroelectric polarization will be along $+x/-x$ and $+y/-y$, respectively. Two purple dots stand for the two ferroelastoelectric ground states at stress-free condition, while the yellow dots stand for the corresponding lattice parameters upon uniaxial tensile strain along $x$ and $y$ axis. The contour colors illustrate the total energy of unit cell relative to the energy of ferroelectric ground states (purple dots). (b) Minimum energy pathway of coherent ferroelastic transition. The gray dots in (b) correspond to the white stars in (a).</td>
<td>38</td>
</tr>
</tbody>
</table>
3.3 Ferroelectric order, spontaneous polarization, and elastic strain effect in monolayer GeSe. (a) Potential energy surface with fractional shift of Se atoms along $x$ and $y$ with the fixed ground-state lattice parameters. Green lines: MEP for ferroelectric transition. Red line: adiabatic path for polarization calculation. Two black dots: degenerate ferroelectric ground states. (b) Double-well potential of monolayer GeSe along the adiabatic path. (c) Calculated total polarization as a function of normalized displacement where the centrosymmetric paraelectric phase (0% displacement) is at the center, and two ferroelectric ground states are at two ends. $P_q$ represents the polarization quanta. (d) Effect of uniaxial strain on the total energy and total spontaneous polarization. (e) Spontaneous polarization of ferroelectric GeSe monolayer as function of uniaxial strain. (f) Minimum energy barrier of coherent ferroelectric transition as function of uniaxial strain without the constraint through the paraelectric phase. 

3.4 Ferroelectric domain wall in monolayer GeS. (a) Ground-state atomistic structure of 180° domain wall. Blue and purple arrows illustrate the polarization direction on the left-hand and right-hand side of the domain wall. (b) MEP of domain wall motion with initial, intermediate, and final configurations labeled from 1 to 9. (c) Relative displacement of each local Ge-S pair along the $x$ direction. Colors from blue to red correspond to the configurations #1-#9 marked in (b). (d) Band structure of monolayer GeS supercell with 180° ferroelectric domain wall. Color indicates the relative localization of electronic states around the domain wall. (e) Top and side view of Kohn-Sham wavefunction at the conduction band minimum (CBM). Purple and green isosurfaces represent the positive and negative values of wavefunctions, respectively.

3.5 Ferroelectric domain wall (180°) in monolayer GeS. (a) Minimum energy pathway of domain wall motion with initial, intermediate, and final configurations labeled from 1 to 9. (b) Atomic structures of the initial (#1), middle (#5), and final (#9) configuration. Blue and purple arrows illustrate the polarization direction of the left-hand and right-hand side of domain wall in monolayer GeS, respectively. (c,d) Relative displacement of each Ge-S pair along $+x$ and $+y$ direction, respectively. Colors from blue to red refer to the configurations #1-#9, corresponding to the colors of the dots marked in (a). (e) 2D plots of relative displacement of each Ge-S pair for the configurations #1, #5, and #9 in the supercell. Small blue and purple arrows illustrate the relative displacement direction of each Ge-S pair. Gray spheres stand for the centers of each Ge-S pair.

3.6 Electronic structures of monolayer GeS with 180° domain wall. (a) Electronic band structure. (b) Top and front view of the Kohn-Sham wavefunction at the conduction band minimum (CBM). (c) Top and front view of the Kohn-Sham wavefunction at the CBM+2. Domain wall is illustrated by dashed blue line. Blue and purple arrows indicate the polarization direction on the left and right hand side of the domain wall.
3.7 Electronic band structures of monolayer GeS, GeSe, SnS, and SnSe. Blue lines: calculations using DFT-PBE with spin-orbit coupling included. Red lines: calculations using many-body perturbation theory within the G$_0$W$_0$ approximation with the spin-orbit coupling included as well. .................................................. 48

3.8 Electronic structure of monolayer GeSe. (a) Band structure of pristine monolayer GeSe. (b) Imaginary dielectric function along the $x$ (red) and $y$ (blue) direction. Two arrows in the band structure (a) indicate the quasiparticle gap that corresponds to two excitonic peaks in the imaginary dielectric function (b). (c) Absorbance (%) of monolayer GeSe. ................................................................. 50

3.9 Mechano-opto-electronic device concepts and their mechanisms. (a) 2D ferroelectric memory. (b) 2D ferroelastic memory. (c) 2D ferroelastoelectric nonvolatile photonic memory. (d) 2D ferroelectric excitonic photovoltaics. .......................... 52

3.10 Atomic configuration of ternary transition metal chalcogenides MM’T$_4$. (a) Top and side view of two-dimensional MM’T$_4$ in monolayer 1T’. Dashed purple lines indicate in-plane zig-zag metallic bond along the $x$ direction owing to Peierls distortion in the $y$ direction. (b) Side view of three-dimensional MM’T$_4$ in their bulk T$_d$ structure. Both M and M’ are transition metal atoms. M, M’, and Te are indicated by blue, orange, and gray atoms. M stands for Ta or Nb, M’ stands for Ir or Rh in the present work.................................................................. 56

3.11 Electronic band structures of 1T’-MM’T$_4$ with and without spin-orbit coupling by DFT calculations with GGA and hybrid HSE06 exchange-correlation functional. 58

3.12 Electronic structure of quantum spin Hall insulators 1T’-MM’T$_4$ using monolayer 1T’-NbRhTe$_4$ as a representative case. (a) 2D Brillouin zone of monolayer 1T’ NbRhTe$_4$ and its projections onto the $k_x$ and $k_y$ edges. Dirac points “Λ” in the absence of spin-orbit coupling are indicated by red dots along the YRY zone boundary. (b) Electronic band structure of monolayer NbRhTe$_4$ with and without spin-orbit coupling indicated by red and blue lines. (c, d) Bulk DOS, edge DOS, and spin ($S_z$)-resolved edge DOS of monolayer NbRhTe$_4$ along the $k_y$ and $k_x$ edge, respectively. Red and blue color in the edge spin DOS indicate the positive and negative spin, respectively. ............................................. 60

3.13 Bulk DOS, edge DOS, and total bulk and edge DOS of four monolayer 1T’-MM’T$_4$. Left panels: along the $Y$. Right panels: along $X$ edges. ......................... 62
3.14 Electronic structure of Weyl semimetal bulk Td-NbRhTe4. (a) 3D Brillouin zone of Weyl semimetal bulk Td-NbRhTe4, and its projections onto top and bottom surfaces along the z direction. Eight gapless Weyl points are located in the \( k_z = 0 \) plane, indicated by red and blue dots (i.e. Chern number of +1 and -1). “W1” and “W2” represent the two Weyl points in the first quadrant of \( k_z = 0 \) plane. Their projections on top and bottom (001) surfaces form Fermi arcs passing through individual pairs of Weyl points. (b,c) Band structure close to “W1” and “W2”, respectively. Conduction and valence bands intersect with a constant energy plane (indicated by a gray mesh), forming a linear cone and separating electron and hole pockets. (d,e) Berry curvature \( \Omega \) around Weyl points \( W_2 \) and \( W_2 \) with high anisotropy, respectively. .......................................................... 64

3.15 Bulk and surface density of states of bulk Td-NbRhTe4. (a,e) Bulk DOS at the two energy levels of 0.125 eV and -0.066 eV corresponding to Weyl points \( W_1 \) and \( W_2 \), respectively. (b,f) Bulk DOS of the local region around the two Weyl points in the first and fourth quadrants (marked by green dashed box in a and e). (c,g) Surface DOS on the top surface of NbRhTe4. (d,h) Surface DOS on the bottom surface of NbRhTe4. Weyl points are marked by red and blue dots on each figure. ............ 65

3.16 Topological phase transition from stacked QSHIs to Weyl semimetal. (a) Topological evolution of Weyl points locked on the \( k_z = 0 \) plane with decreasing effective interlayer distance in bulk Td-NbRhTe4. Solid dots: Chern number of +1; empty dots: Chern number of -1. Type I and Type II Weyl points are labeled at the beginning and the end of the evolution trajectories in the first quadrant. Final eight Weyl points at equilibrium c are marked by dark asterisk symbols “*”. Four groups of Weyl points are indicated by 1-4, respectively. (b-e) Creation, evolution, and annihilation of the first group of Weyl points in Figure 3.15a as interlayer distance decreases captured by a simple \( k \cdot p \) model. Three ellipses correspond to \( E_{1+} \) (blue), \( E_{1-} \) (red) and \( E_2 \) (green), as \( m_{1,2} \) increases. .......................................................... 66

3.17 Dynamic and thermodynamic stability and electronic structure of ternary transition metal tellurides MM’Te4. (a) Phonon dispersion of monolayer 1T’-MM’Te4 using NbRhTe4 as an example. (b) Energy difference between the 1H and 1T’ structure of monolayer MM’Te4. (c) Interlayer binding energies of MM’Te4 and graphite using the optimized optB88 exchange correlation functional for vdW interactions. Graphite is included to serve as a general reference only. (d) Fundamental electronic gaps of monolayer MM’Te4. ......................................................... 71

3.18 Phonon band structures of monolayer 1T’-MM’Te4........................................ 73

4.1 Sheet SHG susceptibilities in GeSe, SnSe, GeS, SnS, MoS2, and h-BN monolayers with their magnitude, real, and imaginary component. ........................................... 80
4.2 Atomic configuration and electronic structures of group IV monochalcogenide MX monolayers. (a) Top and side views of 2D MX. Dashed orange lines indicate the primitive cell with spontaneous polarization along the y-direction. (b) The 2D surface plot of electronic band structure in GeSe monolayer in the first Brillouin zone: CBM (upper) and VBM (lower). \( \Lambda \) denotes the location of energy gap in the reciprocal unit cell, \( \Lambda(k_x, k_y) = [0, \pm 0.39] \) in the fractional coordinate unit. (c) Quasiparticle band structure of GeSe monolayer with the quasiparticle band gap, the lowest exciton, and the exciton binding energy labeled in the plot. ............................. 81

4.3 Sheet SHG susceptibilities in GeSe, MoS\(_2\), and h-BN monolayers with their magnitude, real, and imaginary component. (a) GeSe monolayer which has seven non-zero susceptibility tensor elements with five independent ones due to its point group \( \text{C}_{2v} \) (mm2). (b) MoS\(_2\) monolayer which has four SHG elements with only one independent element due to its point group \( \text{D}_{3h} \) (6m2). (c) h-BN monolayer with only one independent element due to the same point group. Black dots indicate the experimental values. [165] ........................................................................ 83

4.4 Benchmark calculation of the Kramers-Kronig relation. Red and blue lines are the computed imaginary and real component of susceptibility tensor element \( \chi^{(2)}_{yyxx} \) of GeSe monolayer, respectively. Green dots show the real part evaluated using the imaginary part based on the Kramers-Kronig relation, in excellent agreement with the directly-calculated real component (blue lines). ............................. 84

4.5 Polarization anisotropy of sheet SHG susceptibilities \( |\chi^{(2)}_{\parallel}(\theta)|^2 \) and \( |\chi^{(2)}_{\perp}(\theta)|^2 \) in (a) GeSe, (b) MoS\(_2\), and (c) BN marked by red and blue line, respectively. “\( \parallel \)” (“\( \perp \)””) stands for the polarization components of the SHG response parallel (perpendicular) to the polarization \( E(\omega) \) of the incident electric field. “\( \theta \)” is the rotation angle between \( E(\omega) \) and the crystal lattice (i.e. the x axis here for all three 2D materials). Their maximum values are indicated by the green circles in the polar plot and listed at the bottom.............................. 86

4.6 Frequency-dependent (a) real and (b) imaginary part of the interband \( \chi_e(\omega), \chi_e(2\omega) \), intraband \( \chi_i(\omega), \chi_i(2\omega) \), and modulation \( \chi_m(\omega) \) contributions to the sheet SHG susceptibility \( \chi^{(2)}_{yyxx} \). “\( \omega \)” and “\( 2\omega \)” denote the poles of the specific terms. The blue and red dots indicate two major \( 2\omega \) terms from the intraband contribution, \( \chi_i^\parallel(2\omega) \) and \( \chi_i^\perp(2\omega) \). Total real and imaginary part of the SHG susceptibility \( \chi^{(2)}_{yyxx} \) are marked by black line. ............................. 88

xv
4.7 The k-point dependent distribution of sheet SHG susceptibility $\chi_{yxx}^{(2)}$ and $2\omega$ intra-band terms $\chi_{i}^{v}(2\omega)$ and $\chi_{i}^{r}(2\omega)$ in the first BZ of GeSe monolayer. (a, b) Real and imaginary part of $\chi_{yxx}^{(2)}$ dominated by the contributions from two equivalent valleys at the two bandgap locations (marked by $\Lambda$). (c-f) Real and imaginary part of $2\omega$ intraband terms $\chi_{i}^{v}(2\omega)$ and $\chi_{i}^{r}(2\omega)$. (g) The distribution of velocity difference $\Delta_{mn} = v_{mm} - v_{nn}$ between CBM and VBM in the first BZ. (h-i) Real and imaginary part of non velocity terms in $\chi_{i}^{(2)}(2\omega)$. (j, k) Real and imaginary part of $\chi_{i}^{r}(2\omega)$ from three dominating sets of bands.

4.8 Strong coupling among ferroelectric ($P$), ferroelastic ($\eta$) order parameters, and sheet SHG susceptibilities $\chi^{(2)}$. (a) Four different multiferroic states in the GeSe monolayer: $(-P_x, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0)$, $(P_y, \eta_{yy}, \chi_{yxx} \neq 0, \chi_{xyy} = 0)$, $(-P_y, \eta_{yy}, \chi_{yxx} \neq 0, \chi_{xyy} = 0)$, and $(P_x, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0)$ marked by cyan, red, magenta, and green box, respectively. (b-e) Frequency dependent sheet SHG susceptibilities $\chi_{yxx}$ (solid line) and $\chi_{xyy}$ (dashed line) of the above four different multiferroic states in the unit of $10^4 \text{ pm}^2/\text{V}$.

5.1 Microscopic interpretation of shift current and circular current using two-band model. $\delta \langle a \rangle$ is the variation of the mean value of position operator indicating the shift of electron wave packet in real space. Photoexcitation induces the shift of the electron wave packet in real space. SC comes from the displacement of wave packet upon photoabsorption, while CC stems from the asymmetric motion of electrons and holes and the self-rotation of the wave packet. The latter induces itinerant orbital magnetic momentum coupled with the circularly polarized light. Transition rate $W$ of SC is proportional to the linear optical absorption strength under linearly polarized light at frequency $\omega$, while transition rate $F$ of CC is proportional to the local Berry curvature under circularly polarized light at frequency $\omega$.

5.2 Shift current susceptibility tensor elements ($\sigma_{yxx}^{(2)}$, $\sigma_{yxx}^{(2)}$), and circular photocurrent susceptibility tensor elements ($\eta_{yxx}^{(2)}$, $\eta_{yxx}^{(2)}$) of monolayer GeS with and without scissor operator. The scissor operator is based on the bandgap correction using quasiparticle GW gap and exciton binding energy.

5.3 Shift current and circular photocurrent in monolayer MX ($C_{2v}$) and 1H-MX$_2$ ($D_{3h}$) without spin orbit coupling. (A) Shift current susceptibility in 1H-MX$_2$. (B) Shift current susceptibility in MX. (C) Circular photocurrent susceptibility in MX.

5.4 Shift current and circular photocurrent in monolayer MX ($C_{2v}$) and 1H-MX$_2$ ($D_{3h}$) with spin orbit coupling. (A) Shift current susceptibility in 1H-MX$_2$. (B) Shift current susceptibility in MX. (C) Circular photocurrent susceptibility in MX.
5.5 Shift current (SC) and its microscopic origin in 2D ferroelastic-ferroelectric monolayer group IV monochalcogenide GeS. (A) Crystal structure of monolayer group IV monochalcogenides MX where M=(Ge, Sn) and X=(S, Se). (B) 2D electronic band structure near the Fermi level. (C) and (D) Frequency-dependent nonlinear shift current response to incoming linearly $x$ and $y$ polarized light, respectively. (E) and (F) Reciprocal vector $k$ resolved SC susceptibility under linearly $y$ polarized light at the first two peaks (2.0 and 2.8 eV). (G) and (J) $k$-resolved SC strength, (H) and (K) $k$-resolved dipole transition strength, (I) and (L) $k$-resolved topological shift vector of 2D GeS in 2D Brillouin zone under linearly $x$ and $y$ polarized light, respectively. 

5.6 Microscopic distribution and frequency dependent shift photocurrent susceptibility in monolayer 1H-MoSe$_2$ with D$_{3h}$ point group. (A, E) Shift current susceptibility tensor elements $\sigma_{2,xx}$ and $\sigma_{2,yy}$ for shift current using four bands and SOC. (B-D, F-H) the distribution in k-space contributed to the susceptibility tensor element under incident light with different frequencies. 

5.7 Shift current and its microscopic origin in monolayer 1H-MoSe$_2$. (A) 2D electronic band structure near the Fermi level along with its crystal structure. (B) and (E) $k$-resolved shift current strength, (C) and (F) $k$-resolved dipole transition strength, (D) and (G) $k$-resolved topological shift vector of monolayer 1H-MoSe$_2$ in 2D Brillouin zone under linearly $x$ and $y$ polarized light, respectively. 

5.8 Circular photocurrent (CC) and its microscopic origin in 2D ferroelastic-ferroelectric monolayer group IV monochalcogenide GeS. (A) and (B) Two opposite CC susceptibility tensor elements induced by circularly-polarized light. (C) and (D) Evolution of reciprocal vector $k$ resolved CC susceptibility under circularly polarized light at two different frequencies (2.3 and 2.8 eV). (E) and (F) Group velocity difference and Berry curvature between the highest valence band and the lowest conduction band. The white arrows in (E) denote the calculated group velocity difference at specific $k$ point, and the black curves indicate the associated stream lines. 

5.9 Berry curvature distribution in monolayer 1H-MoSe$_2$. 

5.10 SC and CC switching driven by ferroelectricity. The shift vector and Berry curvature are reversed upon ferroelectric transition. 

6.1 Band structure and intraband Berry curvature of bilayer WTe$_2$ using quasiatomic Wannier function with HSE functional ($\lambda=0.4$) and spin-orbit coupling. a band structure using tight-binding approach. b band structure using Kubo formula. It shows that, except some numerical difference, the Berry curvature patterns calculated using the above two methods are very similar.
6.2 Berry curvature dipole of $-i\text{FE}$ trilayer WTe$_2$ with k-point sampling of $600 \times 600 \times 1$ and of $1000 \times 1000 \times 1$. It clearly shows that $600 \times 600 \times 1$ is enough to reach the convergence, which was subsequently used for other Berry curvature, BCD and SD calculations presented in the thesis. .......................................................... 121

6.3 Crystal structure of monolayer, bilayer, and trilayer WTe$_2$. (a) Monolayer 1T’ WTe$_2$ with centrosymmetric $C_{2h}$ point group. It has a mirror plane $M_y$ and a screw rotation symmetry $C_{2y}$, which leads to the inversion symmetry $I = M_y C_{2y}$. (b,c) Bilayer and trilayer Td WTe$_2$ with $C_s$ point group. $C_{2y}$ symmetry is broken, hence the inversion symmetry $I$ is also broken in bilayer and trilayer WTe$_2$. .......... 122

6.4 Ferroelectric transition in bilayer and trilayer WTe$_2$. (a-c) ferroelectric transition in bilayer WTe$_2$. The two opposite ferroelectric states ($-m\text{FE}$ and $+m\text{FE}$) in bilayer are transformed through a glide plane operation $\{M_z t_a\}$, that is, a mirror operation $M_z$ followed by an in-plane shift along $x$ by $t_a$ with $t_a = 1/2a + d_x$. The intermediate PE state $C_{2v}$ point group, thus its out-of-plane polarization with vanishes due to the glide plane $\{M_z 1/2a\}$ symmetry. (d-f) ferroelectric transition in trilayer WTe$_2$. The two opposite ferroelectric states ($-i\text{FE}$ and $+i\text{FE}$) in trilayer are related to each other through an inversion operation $\{I0\}$. The intermediate PE state of trilayer WTe$_2$ has $C_{2h}$ point group, thus its out-of-plane polarization also vanishes due to inversion symmetry. The red and green vertical dashed lines show the relative shift $\pm d_x$ between adjacent WTe$_2$ layers. The corresponding in-plane shift is very small ($d_x \sim 20$ pm), therefore it is exaggerated in the above plots for illustrative purpose. .......................................................... 123

6.5 Electric polarization of bilayer WTe$_2$. a Out-of-plane polarization of bilayer WTe$_2$ as a function of distance away from equilibrium configuration, $\Delta d_z$, between two layers. b Total energy as a function of the out-of-plane polarization. ............ 125

6.6 Electronic structure and NAHE of bilayer WTe$_2$. (a) Band structure of bilayer WTe$_2$ in $\pm m\text{FE}$ state color-coded by the z-component of intraband Berry curvature $\Omega^z_n(k)$. Spin-orbit coupling is included and hybrid HSE06 functional is employed. (b) BCD and SD tensor elements $D_{yz}^{\text{intra}}(\mu)$, $D_{yz}^{\text{inter}}(\mu,\omega)$, and $D_{xy}^{\text{inter}}(\mu,\omega)$ as function of chemical potential $\mu$. For interband BCD and SD, $\omega$ is set to 120 meV. (c,d) k-dependent distribution of intraband Berry curvature $\Omega^z_{\text{occ,intra}}(k)$ at $\mu = \pm 50$ meV, respectively. e,f k-dependent distribution interband Berry curvature $\Omega^z_{nm,\text{inter}}(k)$ for $(n,m)$=(VBM-1,CBM) and for $(n,m)$=(VBM-1,CBM), respectively. The results are the same for both $+m\text{FE}$ and $-m\text{FE}$ state, suggesting that nonlinear anomalous Hall current can be induced in bilayer WTe$_2$, but it will not switch sign upon ferroelectric transition. .......................................................... 127
6.7 Electronic structure and FNAHE of trilayer WTe$_2$. (a,b) Band structure of trilayer WTe$_2$ in -iFE and +iFE state, respectively. Both are color-coded by the $z$-component of intraband Berry curvature $\Omega^z_n(k)$. (c,d) BCD tensor elements $D^{\text{intra}}_{yz}(\mu)$ and $D^{\text{inter}}_{yz}(\mu,\omega)$, and SD tensor element $D^{\text{L,inter}}_{xy}(\mu,\omega)$ as function of chemical potential $\mu$ for -iFE and +iFE state, respectively. For interband BCD and SD, $\omega$ is set to 120 meV. (e,f) k-dependent distribution of intraband Berry curvature $\Omega^z_{\text{intra}}(k)$ at $\mu = \pm 50$ meV for -iFE and +iFE state, respectively. (g,h) k-dependent distribution interband Berry curvature $\Omega^z_{\text{inter}}(k)$ between (VBM-1,CBM) around the Fermi surface for -iFE and +iFE state, respectively. The results clearly show that nonlinear anomalous Hall current in trilayer WTe$_2$ will switch sign upon ferroelectric transition, in direct contrast to the bilayer case.

6.8 Interband LPGE susceptibility $\sigma_{abc}$ of bilayer and trilayer WTe$_2$ at $\mu=0$. (a) $\sigma_{abc}$ for the $\pm m$FE states of bilayer WTe$_2$. (b) $\sigma_{abc}$ for the $\pm i$FE state of trilayer WTe$_2$. These two plots demonstrate that, under ferroelectric transition, the sign of LPGE-induced dc current susceptibility switches in the trilayer case, but remains invariant in the bilayer case.

6.9 Band structure of bilayer WTe$_2$ color-coded by spin polarization $S^x_n(k)$ using the PBE functional. (a) band structure for the $-m$FE state. (b) band structure for $+m$FE state. Spin-orbit coupling is included. $S^x_n(k)_{-m\text{FE}} = -S^x_n(k)_{+m\text{FE}}$ holds for arbitrary k point.

6.10 Band structure of bilayer WTe$_2$ color-coded by spin polarization $S^z_n(k)$ using the PBE functional. (a) band structure for the $-m$FE state. (b) band structure for $+m$FE state. Spin-orbit coupling is included. $S^z_n(k)_{-m\text{FE}} = S^z_n(k)_{+m\text{FE}}$ holds for arbitrary k point.

6.11 Band structure of trilayer WTe$_2$ color-coded by spin polarization $S^z_n(k)$ using the PBE functional. (a) band structure for the $-i$FE state. (b) band structure for $+i$FE state. Spin-orbit coupling is included. $S^z_n(k)_{-i\text{FE}} = -S^z_n(k)_{+i\text{FE}}$ holds for arbitrary k point.
6.12 Transformation of pseudovectors and polar vectors under different symmetry operation and ferroelectric switching of LPGE and CPGE nonlinear current. (a) Transformation of Berry curvature, spin polarization, shift vector, BCD and SD as well as LPGE and CPGE nonlinear current under different symmetry operation between two ferroelectric states in time-reversal invariant few-layer WTe$_2$. Berry curvature and spin polarization transform as time-reversal antisymmetric pseudovectors. Under a mirror symmetry operation $M_z$ for the ferroelectric states in bilayer WTe$_2$, most quantities remain invariant, except in-plane spin and Berry curvature component. Under inversion symmetry operation $I$ for the $-\text{iFE}$ and $+\text{iFE}$ state states in trilayer WTe$_2$, all quantities, including Berry curvature, spin polarization, shift vector, BCD and SD, flip the sign in the presence of time-reversal symmetry, giving rise to FNAHE in trilayer WTe$_2$. (b,c) Ferroelectric switching of nonlinear current in the $-\text{iFE}$ and $+\text{iFE}$ state of trilayer WTe$_2$, respectively. Upon the polarization $P_z$ switching, nonlinear anomalous Hall current $j^L_x$ from LPGE switches between $-x$ and $+x$ direction under external field with in-plane linear polarization, while nonlinear anomalous Hall current $j^C_y$ from CPGE switches between $-y$ and $+y$ direction under circularly-polarized light with normal incidence. 

7.1 Slab model of 1L, 2L, 7L anatase TiO$_2$ (101) surface. 

7.2 Imaginary and real part SHG in 7L TiO$_2$ and bulk GaAs. The effective thickness of layered TiO$_2$ included ~15 Å of vacuum. The SHG susceptibility component in GaAs is $\chi^{xyz}$. 

7.3 Interband dipole transition strength in the first Brillouin zone. VBM denotes valence band maximum and CBM represents conduction band minimum. 

7.4 Atomic structure of bilayer antiferromagnetic MnBi$_2$Te$_4$. The magnetic orderings on Mn atoms are marked by blue (down) and red (up) arrows. The right part is the band structure and band energy difference in the Brillouin zone between VBM and CBM. 

7.5 The distribution of $\Delta x_{ic} \{r^x_{ic}, r^x_{vc}\} \delta (\omega_{mn} - \omega)$ in Brillouin zone at $\omega=0.12$ eV and $\omega=0.60$ eV in AFM bilayer MnBi$_2$Ti$_4$. 

7.6 The distribution of $\Delta x_{ic} \{r^x_{ic}, r^x_{vc}\} \delta (\omega_{mn} - \omega)$ in Brillouin zone at $\omega=0.12$ eV and $\omega=0.60$ eV in $T$ AFM bilayer MnBi$_2$Ti$_4$. 

7.7 Shift current in bilayer AFM MnBi$_2$Te$_4$ under electric field $E_z=\pm 0.01$V/Å and convergency of k-point sampling in the tight-binding calculation. 

7.8 The distribution of $\sum_{nm\sigma} f_{nm} R_{nm}(k) f^y_{nm} r^y_{mn} \delta (\omega_{mn} - \omega)$ in the first Brillouin zone at $\omega=0.36$ eV and $\omega=0.60$ eV in AFM bilayer MnBi$_2$Ti$_4$. 

xx
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Second order nonlinear photocurrent responses under different polarized light. ↔ and ↓ indicate horizontal (x) and vertical (y) linearly-polarized light, respectively. ⊙ and ⊚ indicate left and right circularly polarized light, respectively. $E_0$ is the strength of the applied electric field. $\varphi$ stands for linearly-polarized light that has a 45 degree angle with respect to $x$ and $y$. Shift current (SC) and circular photocurrent (CC) can coexist in 2D MX (with C$_{2v}$ point group) upon left/right circularly polarized light. However, only SC will persist, while CC will disappear in 2D MX under linearly polarized light such as ↔, ↓, and $\varphi$. It is important to mention that both SC and CC in monolayer MX$_2$ (e.g. 1H-MoSe$<em>2$ with D$</em>{3h}$ point group) will be always zero under left/right circularly polarized light.</td>
<td>107</td>
</tr>
<tr>
<td>5.2 Transformation of inter-band Berry connections $r_{mn}$, shift vector $R_{mn}$, group velocity difference $\Delta_{mn}$ and Berry curvature $\Omega_{mn}$ under space inversion $I$ and time reversal $T$ symmetry operation. $R_{mn}(k)$ is odd under $I$ and even under $T$ in momentum space. $\Omega_{mn}(k)$ is even under $I$ and odd under $T$ in momentum space. These transformation rules govern the coupling between ferroelectric polarization and nonlinear SC and CC photocurrent.</td>
<td>110</td>
</tr>
<tr>
<td>6.1 The correlation between the irreducible representations of parent group C$_{2h}$ and its noncentrosymmetric subgroups C$_2$, C$_s$, and C$_1$.</td>
<td>132</td>
</tr>
</tbody>
</table>
1. INTRODUCTION TO NONLINEAR OPTICS

1.1 The Birth of Nonlinear Optics

Light has the characteristic properties of frequency, amplitude, phase, and spin and orbital angular momentum, which can be used to probe and control the physical properties of matter. Light-matter interaction thus has deep impact in multiple disciplines such as atomic, molecular, and optical physics (AMO), condensed matter physics, materials science, electrical engineering, and now increasingly in biology and medicine. As in the words of Niels Bohr, “Isolated material particles are abstractions, their properties being definable and observable only through their interactions with other systems.” The processes are the results of the interaction between electromagnetic field and charged particles which can be observed either by the changes to the light, e.g., linear optical absorption, or by the changes to the matter, e.g., ionization. As a matter of fact, the world we are living is quite nonlinear with the excitement of nonlinearity that linear physical law could not furnish. There has been tremendous interest in nonlinear optics which considers the consequence of the modification of the linear optical properties of a material system in the presence of light [1]. Many unique optical phenomena have been brought about by the availability of high-intensity laser beams. For example, Franken et al. used the ruby laser for the second-harmonic generation (SHG) in 1961 [2], which is marked as the beginning of nonlinear optics. More interestingly, the ruby laser itself is a paradigm example of nonlinear optics.

1.2 What Is Nonlinear Optics?

The classical electrodynamics tell us that photons do not interact with each other. This is not true from the point view of quantum electrodynamics: a photon can fluctuate into electron-positron pair which then couples to another photon by the higher order process, e.g., Delbrück scattering, as illustrated in Figure 1.1. The photon-photon scattering has never been discovered in vacuum in the field of two-photon physics (also called gamma-gamma physics) so far. However, this light-by-light interaction readily and commonly takes place in material systems corresponding to nonlinear
optical (NLO) responses. Nonlinear optics is the subject of the interaction of intense laser field with matter, and the laser made NLO phenomena possible by its sufficiently intense field, which can modify the optical properties of a material. NLO phenomena are “nonlinear”, in the sense that the responses in polarization or magnetization depend on the optical field in a nonlinear way (e.g., second-order, third-order polynomial).

\[ \chi_1 e^{-\chi_3 e^{-\chi_2 e_+}} \gamma_4 \]

Figure 1.1: A Feynman diagram for photon-photon scattering $\gamma_1 \gamma_2 \rightarrow \gamma_3 \gamma_4$ with electron-positron pair $e^- e^+$

The mathematical formulas of the electric nonlinearities can be derived by considering how polarization $P(t)$ and magnetization $M(t)$ depend on the electric part of the applied optical field $E(t)$. Upon external electromagnetic (EM) field, $E(t) = \sum_p E(\omega_p) e^{-i\omega_p t}$. The induced electric dipole $P(t) = \sum_p P(\omega_p) e^{-i\omega_p t}$ can be expanded in linear and higher orders of $E$:

\[
P = P^{(0)} + P^{(1)} + P^{(2)} + P^{(3)} + \ldots, \tag{1.1}
\]

where

\[
P^{(1)}(\omega_p) = \chi^{(1)}(-\omega_p; \omega_p) : E(\omega_p)
\]

\[
P^{(2)}(\omega_p + \omega_q) = \chi^{(2)}(-\omega_p - \omega_q; \omega_p, \omega_q) : E(\omega_p) E(\omega_q)
\]

\[
P^{(3)}(\omega_p + \omega_q + \omega_r) = \chi^{(3)}(-\omega_p - \omega_q - \omega_r; \omega_p, \omega_q, \omega_r) : E(\omega_p) E(\omega_q) E(\omega_r).
\]
\( \chi^{(1)}, \chi^{(2)}, \chi^{(3)} \) are the linear, second-order, and third-order optical susceptibilities in Gaussian units, respectively. The permittivity or dielectric constant is directly related to the linear optical properties of materials and the relation between dielectric constant and the linear susceptibility is given by

\[
\varepsilon = 1 + 4\pi \chi = \varepsilon_1 + i\varepsilon_2. \tag{1.3}
\]

The complex index of refraction of the material system is defined as \( N = \sqrt\varepsilon = n + ik \) [3], where \( n \) and \( k \) are the refractive index and extinction coefficient given by

\[
n = \left[ \frac{\sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_1}}{2} \right]^{\frac{1}{2}} \tag{1.4}
\]

\[
k = \left[ \frac{\sqrt{\varepsilon_1^2 + \varepsilon_2^2 - \varepsilon_1}}{2} \right]^{\frac{1}{2}}.
\]

The absorption \( \alpha \), energy-loss spectrum \( L \), and reflectivity \( R \) can be directly measured in experiments, which can be written as

\[
\alpha = \frac{2\omega}{c} k
\]

\[
L = \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2} \tag{1.5}
\]

\[
R = \frac{|1-N|^2}{|1+N|} = \frac{(1-n)^2 + k^2}{(1+n)^2 + k^2}.
\]

It should be noted that nonlinear polarization \( \mathbf{P}^{NL} \) would act as an inhomogeneous source term in the Maxwell equation [4]

\[
\nabla^2 \mathbf{E} - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{4\pi}{c^2} \frac{\partial^2 \mathbf{P}^{NL}}{\partial t^2}, \tag{1.6}
\]

which will play a significant role in the description of EM wave propagation in the media. More particularly, the nonlinear terms can be comparable to the linear term when the amplitude of the applied electric field \( \mathbf{E} \) is sufficient intense as mentioned above. Here we can do a rough numerical estimation based on the characteristic atomic electric field strength \( E_{a_0} = \frac{e}{a_0^2} = 2 \times 10^7 \) statvolt/cm where \( E \) becomes comparable to the binding electrons in atom. It gives a laser intensity of \( I_{a_0} = \frac{e}{8\pi} E_{a_0}^2 = 5 \times 10^{23} \) erg/cm\(^2\)s = 5 \times 10^{16} \) W/cm\(^2\), which is possible to achieve with laser techniques. In fact, the record for intensity of a laser beam is about \( 10^{23} \) W/cm\(^2\) corresponding to.
spotlighting all the sunlight power that bombard the Earth onto a single sand grain.

Many NLO phenomena with distinct features have been encoded in the nonlinear susceptibilities. Hence, our goal is to develop the microscopic theory and understanding of the nonlinear susceptibilities and enable predictive discovery and design of NLO materials. We illustrate some typical 2\textsuperscript{nd} and 3\textsuperscript{rd} NLO responses in Figure 1.2, e.g., second harmonic generation, linear electro-optic effect, static polarization (current, current injection), third harmonic generation, quadratic electro-optic effect, and two-photon absorption with a two-band energy-level description based on equation (1.2). For example, in second harmonic generation, a new photon (denoted by blue wave arrow) is generated with doubled frequency of the initial photons (denoted by red wave arrows) and the induced electric polarization depends quadratically on the strength of the applied electric field. We shall discuss the general theory, discovery, and potential applications of those NLO effects in the following chapters.

Figure 1.2: Typical NLO effects: second harmonic generation, linear electro-optic effect, static polarization (current, current injection), third harmonic generation, quadratic electro-optic effect, and two-photon absorption. \( E \) is a static electric field.
1.3 Motivation

NLO has rich fundamental physics and has led to many important functional devices in lasers, frequency conversion module, electro-optic modulators, switches, and so forth. In particular, those devices are highly desirable for many scientific disciplines and technological applications [5], e.g., ultrafast nonlinear optics [6], nonlinear biosensing and imaging [7], efficient generation of entangled photon pairs for quantum computing and quantum sensing [8,9], and all-optical transistor and computer [10–12].

Quantum entanglement is crucial for fundamental tests of nonlocality in quantum mechanics and for secure communications including quantum teleportation and quantum cryptography which are theoretically unhackable [13]. As one example of the NLO applications, the most extensively used method for generating entangled photon pairs is spontaneous parametric down conversion (SPDC) [14,15], which is a reverse process of nonlinear optical sum-frequency generation (SFG) or SHG characterized by $\chi^{(2)}(-\omega_1, -\omega_2; \omega_3)$ with energy conservation of $\hbar \omega_3 = \hbar \omega_2 + \hbar \omega_1$. The signal and idler photons can be entangled in polarization [16]

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( |H_1 H_2\rangle + e^{i\phi(\omega_p, \omega_s, \omega_i, k_p, k_s, k_i)} |V_1 V_2\rangle \right), \quad (1.7)$$

where $|H\rangle$ and $|V\rangle$ denote horizontal and vertical polarized photons. $\omega_p, \omega_s, \omega_i$ and $k_p, k_s, k_i$ are frequencies and momentum of the pump, signal, and idler, respectively. The entangled photon pair has no specific polarization in this superposition state until the polarization of either photon is observed, consequently, the polarization of the other photon is immediately determined, regardless of how far away they have been separated.

In non-parametric NLO processes, energy, momentum, or angular momentum can be transferred to modify the quantum states of the medium. For example, the non-parametric NLO processes can electromagnetically induce transparency [17,18]. Based on this process, the light speed was reduced to 17 m/s in an ultra-cold atomic gas with reflective index modified by exceptionally large optical nonlinearities [19]. Two-photon absorption (TPA) [20] occurs when two photons are
simultaneously absorbed in the media, which is also a non-parametric NLO process and widely used in fluorescence microscopy in the near-infrared region.

NLO process also plays an important role in bulk photovoltaic effect (BPVE) in noncentrosymmetric materials [21], offering alternative approach for photocurrent generation beyond conventional $p–n$ junction based solar cells whose performance is limited by the Shockley-Queisser limit [22]. More importantly, the BPVE allows to generate a photovoltage far beyond the bandgap which paves way for new generation photovoltaics, e.g., shift current response [23].

Moreover, symmetry plays a crucial role in the discovery of the fundamental laws in physics in last century and continuously serves as an overarching principle in seeking further progress [24]. The spontaneous symmetry breaking gives rise to various phenomena related to physical observables, which can be understood in terms of symmetry-operational similarities (SOS) [25]. However, there are many symmetry-related physical phenomena that are challenging to observe. NLO responses render the possibility to serve as an effective and powerful tool to probe various symmetry-related quantities. For example, one can use SHG to observe the subtle antiferromagnetic ordering in layered chromium triiodide (CrI$_3$) [26]. The nonvanishing NLO responses are immediately determined by the point group symmetry including magnetic symmetry. Topology, another critical theme in condensed matter physics, becomes a blossoming field with many challenging fundamental questions in terms of many-body interaction, crystalline symmetries, coupling to the external fields, etc. Again, NLO responses can help reveal the topological properties of materials. For example, an exotic Berry curvature dipole induced nonlinear Hall effect was predicted in time-reversal invariant materials [27, 28]. In addition, the nonlinear circular photogalvanic effect (CPGE) enables the direct observation of the chiral monopole charge in Weyl semimetals [29].

Together with symmetry and topology, NLO leads to rich extraordinary physical phenomena in various phases of matter. As a consequence, NLO enables many important applications and powerful experimental tools based on the nonlinear coupling between optical field and electronic states of matter.
This chapter will discuss the theoretical frameworks for NLO responses. A general formulation of electric (magnetic) nonlinearities is based on a series expansion of the electric polarization or magnetization in powers of the amplitude of electric (magnetic) field as previously mentioned and described by Eq. (1.2). Here we present a perturbation theory and obtain the quantum mechanical expressions of nonlinear susceptibilities that are amenable for numerical calculations using first-principles theory and tight-binding approach. The symmetry of nonlinear susceptibilities is fully addressed within group theory.

2.1 A Classical Model of Second-Order Electrical Nonlinearities

First, let us consider a classical modified harmonic oscillator with an anharmonic potential, known as Lorentz model

\[ V(x) = \frac{1}{2} m \omega_0^2 x^2 + \frac{1}{3} m \kappa x^3, \tag{2.1} \]

where \( \kappa \) represents the strength of the nonlinearity, \( m \) is the mass of electron, and \( \omega_0 \) is the harmonic oscillating frequency, and \( x \) is the electron coordinate. The gradient of potential gives a restoring force

\[ F_{\text{restoring}}(x) = -\nabla V(x) = -m \omega_0^2 x - m \kappa x^2. \tag{2.2} \]

Under external electric field \( E(t) = E_0 e^{i\omega t} \), the equation of motion can be written as

\[ m \frac{d^2 x}{dt^2} + m \gamma \frac{dx}{dt} + m \omega_0^2 x + m \kappa x^2 = -e E(t), \tag{2.3} \]

where \( \gamma \) is the damping parameter due to energy-dissipating forces. Using a trial solution with plane wave form \( x(t) = x_1 \exp(i \omega t) + x_2 \exp(i 2 \omega t) \), we can obtain the solution for the above
equation of motion

\[ x_1 = -\frac{eE_0}{m} \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega} \]
\[ x_2 = -\kappa x_1^2 \frac{1}{\omega_0^2 - (2\omega)^2 + i\gamma 2\omega}. \]  

(2.4)

The electronic polarization is given by

\[ P = -N e x = \chi^{(1)} E_0 + \chi^{(2)} E_0^2. \]

Hence, we get the linear and second order nonlinear susceptibilities

\[ \chi^{(1)}(\omega) = \frac{Ne^2}{m} \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega} \]
\[ \chi^{(2)}(2\omega) = \kappa \frac{Ne^3}{m^2} \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega} \frac{1}{\omega_0^2 - (2\omega)^2 + i\gamma 2\omega}. \]  

(2.5)

In a centrosymmetric media, where \( V(x) = V(-x) \Rightarrow \kappa = 0 \), the second-order nonlinear susceptibility is zero. While this toy model provides a simple picture of electric nonlinearities, to get more fundamental insights of the NLO responses, we will apply the quantum theory as detailed next.

2.2 Density Matrix Formalism and Time Dependent Perturbation Theory

In quantum mechanics, it is often beneficial to avoid the coordinate representation in many-particle Hamiltonian by transforming from Hilbert space to Fock space where the physical states of a system of \( N \) electrons are vectors spanning \( N \)-particle subspace. A second-quantization formalism in the occupation number representation describes systems of many particles more generally and yields many-particle Hamiltonian with the same form as the single-particle Hamiltonian.

2.2.1 The Interaction Hamiltonian in Second Quantization

We start with a many-particle Hamiltonian including spin-orbit coupling (SOC) and the remaining many-body interactions \( H_{\text{rest}} \), e.g., electron-electron interaction, electron-phonon coupling

\[ H_0 = \int d\mathbf{r} \Psi^\dagger \mathbf{H}_o \Psi + H_{\text{rest}}, \]

(2.6)

where

\[ \mathbf{H}_o = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) + \frac{e\hbar}{4m^2c^2} [\nabla V(\mathbf{r}) \times \mathbf{p}] \cdot \mathbf{\sigma} \]  

(2.7)
is the ground state Hamiltonian density, $V(r)$ is the periodic lattice potential and $\sigma$ is the Pauli spin matrix. The field operator $\Psi = \begin{pmatrix} \psi^\uparrow \\ \psi^\downarrow \end{pmatrix}$ is spin resolved and satisfies the anticommutation relation $\{\psi^\uparrow(r,t), \psi^\dagger_\downarrow(r',t)\} = 1_{2\times2} \otimes \delta(r-r')$. A macroscopic Maxwell electric field $E(t)$ can be incorporated in the above Hamiltonian by a minimal substitution of the conjugate momentum $p \mapsto p - eA$, where the electron charge $e = -|e|$ includes the negative sign. The relation between electric and magnetic fields $E(t), B(t)$ and vector potential $A(t)$ reads

$$E(t) = -\nabla \phi - \frac{1}{c} \frac{\partial A(t)}{\partial t},$$

$$B(t) = \nabla \times A(t).$$

(2.8)

In velocity gauge, we set the scalar potential $\phi = 0$. For example, the circularly polarized light gives the vector potential $A(t) = iE_\omega e^{-i\omega t} - iE_\omega^\ast e^{i\omega t}$. The field operator evolves according to Heisenberg equation of motion

$$\frac{d\Psi(r,t)}{dt} = \frac{1}{i\hbar} [\Psi(r,t), H(t)].$$

(2.9)

We define a new field operator by a gauge transformation $\tilde{\Psi} = \Psi e^{-i\mathbf{A}\cdot\mathbf{r}/(\hbar c)}$, and then arrive the interaction Hamiltonian [30]

$$H_{\text{int}}(t) = \int \tilde{\Psi}^\dagger(r,t) [\mathcal{H}_o - e\mathbf{r} \cdot \mathbf{E}(t)] \tilde{\Psi}(r,t) \, dr + H_{\text{rest}}.$$

(2.10)

It should be noted that the interaction Hamiltonian in velocity gauge is equivalent to the length gauge formalism [31]. The ground state Schrödinger equation is written by

$$\mathcal{H}_o \psi_{n\alpha}(\mathbf{k}; \mathbf{r}) = \hbar \omega_{n\alpha}(\mathbf{k}) \psi_{n\alpha}(\mathbf{k}; \mathbf{r}),$$

(2.11)

where $\alpha$ is the spinor index. For simplicity, we contract the spinor indices into band indices $n \mapsto n\alpha$. The eigenstates satisfy Bloch’s theorem $\psi_n(\mathbf{k}; \mathbf{r}) = u_n(\mathbf{k}; \mathbf{r}) e^{ik\cdot r}$ due to the lattice translation symmetry. $u_n(\mathbf{k}; \mathbf{r})$ is the periodic part of the Bloch function. The field operator can be expanded
in the Bloch functions
\[ \tilde{\psi}(r, t) = \sum_n \int dka_n(k)\psi_n(k; r), \] (2.12)
where the fermion creation and annihilation operators fulfill \( \{a_n(k), a_m^+(k')\} = \delta_{nm}\delta(k - k'). \)

Here we denote \( \int d k \mapsto \int d^l k / (2\pi)^l \) for \( l \)-dimensional system and take the integral of fermion operators \( \langle \Psi | a_n^+(k) a_n(k') | \psi \rangle = V f_n \delta_{kk'} \). For practical numerical integration, we substitute the integral by a direct sum: \( \int d k / (2\pi)^l = 1 / V \sum_k \), where is \( V \) the volume of the crystal. Finally, we obtain the unperturbed Hamiltonian
\[ H_0 = \sum_n \int d k h\omega_n(k) a_n^+(k) a_n(k). \] (2.13)

To deal with the interaction Hamiltonian \( H_{\text{int}}(t) \) in Eq. (2.10), one has to determine the dipole matrix element
\[ \langle n k | r | m k' \rangle \equiv \int \psi_n^*(k; r) r \psi_m(k; r) dr. \] (2.14)

Blount first introduced the renowned identity to separate the dipole matrix elements into intraband and interband terms according to [32]
\[ \langle n k | r | m k' \rangle = \delta(k - k') \xi_{nm} + i\delta_{nm} \nabla_k \delta(k - k'), \] (2.15)
and the intra position operator \( r_i \) and inter position operator \( r_e \) are given by
\[ \langle n k | r_i | m k' \rangle = \delta_{nm} [\delta(k - k') \xi_{nn} + i \nabla_k \delta(k - k')], \] (2.16)
\[ \langle n k | r_e | m k' \rangle = (1 - \delta_{nm}) \delta(k - k') \xi_{nm}, \]
where the Berry connections \( \xi_{nm}(k) \) are defined by the periodic part of the Bloch functions
\[ i \frac{\partial u_n(k; r)}{\partial k} = \sum_m u_m(k; r) \xi_{mn}(k). \] (2.17)

10
The intraband term leads to the troublesome nature of the position operator in periodic systems due to the singularity in $\nabla_k \delta (k - k')$. The derivation is shown in APPENDIX A. We use the same notation in Sipe’s paper for the dipole matrix [30, 31]

$$r_{nm} = \xi_{nm} \quad n \neq m \quad \text{otherwise}.$$  \hspace{1cm} (2.18)

The intraband Berry connection contributes to the modern theory of polarization in crystalline solids and has a geometrical meaning of “connection” of the quantum state manifolds in Hilbert space. Following the approach by King-Smith and Vanderbilt [33], we reformulate the numerical expression of Berry connection

$$A_n^a(k) \equiv \xi_{nm}^a(k) = i \langle nk | \frac{\partial}{\partial k_a} | nk \rangle = i \langle nk | nk' \rangle \frac{\partial}{\partial k_a} \ln \langle nk | nk' \rangle$$

$$= i \lim_{\Delta k_a \to 0} \frac{1}{\Delta k_a} \left( \ln \langle nk | n (k + \Delta k_a q) \rangle - \ln \langle nk | nk \rangle \right)$$

$$\quad = i \lim_{\Delta k_a \to 0} \frac{1}{\Delta k_a} \ln \langle nk | n (k + \Delta k_a q) \rangle.$$  \hspace{1cm} (2.19)

Here, the Berry connection $A_n$ is related to the overlap of the two neighboring Bloch wavefunctions in momentum space as

$$\langle u_n(k) | u_n(k + \delta k) \rangle = e^{i \delta k \cdot A_n(k)}.$$  \hspace{1cm} (2.20)

The curl of intraband Berry connection gives the important topological quantity of Berry curvature, $\Omega_n$, which is directly related the anomalous Hall effect. The definition of Berry curvature $\Omega_n$ is

$$\Omega_n \equiv \nabla \times \xi_{nm} = \nabla \times A_n.$$  \hspace{1cm} (2.21)

It should be noted that the Berry curvature is gauge invariant although the Berry connection is gauge dependent. After substituting the field operator with dipole matrix elements into Eq. (2.10),
we obtain the interaction Hamiltonian in second quantization

\[
H_{\text{int}}(t) = \sum_n \int dk h \omega_n(k) a_n^\dagger(k) a_n(k) \\
- eE(t) \cdot \sum_n \int dk r_{nn}(k) a_n^\dagger(k) a_n(k) \\
- eE(t) \cdot \sum_{n,m} \int dk r_{nm}(k) a_n^\dagger(k) a_m(k) + H_{\text{rest}},
\]

(2.22)

where \( r_{nn}(k) \) is the covariant derivative

\[
r_{nn}(k) \equiv \xi_{nn}(k) - i \nabla k.
\]

(2.23)

One can verify that the interaction Hamiltonian \( H_{\text{int}}(t) \) is gauge invariant under a change in the phases of the Bloch wavefunctions.

2.2.2 The Polarization and Current Density Operator

Eq. (2.22) gives the definition of intraband and interband polarization by \( H = H_0 - V \cdot E + H_{\text{rest}} \), written as

\[
P_{\text{intra}}(t) = \frac{e}{V} \int dk \sum_n r_{nn}(k) a_n^\dagger(k) a_n(k)
\]

(2.24)

\[
P_{\text{inter}}(t) = \frac{e}{V} \int dk \sum_{n,m} r_{nm}(k) a_n^\dagger(k) a_m(k).
\]

The current density operator in the Heisenberg picture reads

\[
J(t) = \frac{e}{mV} \int d\mathbf{r} \psi^\dagger(\mathbf{r}, t) \left[ \frac{\hbar}{i} \nabla - \frac{e}{c} \mathbf{A}(t) + \frac{\hbar}{4mc^2} \mathbf{\sigma} \times \nabla V(\mathbf{r}) \right] \psi(\mathbf{r}, t)
\]

\[
= \frac{e}{mV} \int d\mathbf{r} \tilde{\psi}^\dagger(\mathbf{r}, t) \left[ \frac{\hbar}{i} \nabla + \frac{\hbar}{4mc^2} \mathbf{\sigma} \times \nabla V(\mathbf{r}) \right] \tilde{\psi}(\mathbf{r}, t)
\]

(2.25)

\[
= \frac{e}{V} \sum_{n,m} \int dk v_{nm}(k) a_n^\dagger(k) a_m(k),
\]

where velocity matrix elements are renormalized by spin-orbit coupling and can be expressed by

\[
v_{nm} \equiv \langle n|k|v|m \rangle = i r_{nm} \omega_{nm}
\]

for nondegenerate k-points in the Brillouin zone. Moreover, the
current operator is related to the polarization operator as follows

\[ J(t) = \frac{dP(t)}{dt}. \] (2.26)

Furthermore, we have the equation of motion

\[ \frac{dP(t)}{dt} = \frac{1}{i\hbar} [P, H_{\text{int}}]. \] (2.27)

The total current can be separated into the intraband and interband parts. To calculate the intraband current, we first calculate the interband current without many-body interactions, say \( H_{\text{rest}} = 0 \), from

\[ \frac{dP_{\text{inter}}(t)}{dt} = \frac{1}{i\hbar} [P_{\text{inter}}, H_{0} + \mathcal{V}(P_{\text{inter}} + P_{\text{intra}}) \cdot E]. \] (2.28)

We denote the inter and intra polarization operator by \( P_{e} \) and \( P_{i} \), respectively. The corresponding commutative relation between the different components is given by

\[ [P_{e}, P_{i} + P_{e}] = \frac{-ie^{2}}{\mathcal{V}} \sum_{nm} \int d\mathbf{k} \left( r^{a}_{nm;b} + i \sum_{p} \left[ r^{a}_{np}r^{b}_{pm} - r^{b}_{np}r^{a}_{pm} \right] \right) a_{n}^{\dagger}a_{m}, \] (2.29)

where

\[ r^{a}_{nm;b} \equiv \left[ \frac{\partial}{\partial k_{a}} - i \left( \xi^{a}_{nm} - \xi^{a}_{nm} \right) \right] r_{nm}, \] (2.30)

is the generalized or covariant derivative. The detailed derivation is shown in the APPENDIX A. To simplify the expression of current operator, we apply the sum rule derived from the commutative identity of position operator

\[ [\mathbf{r}^{a}, \mathbf{r}^{b}] = 0. \] (2.31)

Similarly, by decomposing the position operator into intraband and interband parts, one can obtain the sum rule

\[ r^{a}_{nm;b} - r^{b}_{nm;a} = -i \sum_{l} \left[ r^{a}_{nl}r^{b}_{lm} - r^{b}_{nl}r^{a}_{lm} \right]. \] (2.32)
Substituting the sum rule and Eq. (2.29) into Eq. (2.25), we obtain an important result for the intraband current which is valid for both metals and insulators, as reformulated by Fregoso [34]

\[ J_i = \frac{e}{V} \sum_{nm} \int d\mathbf{k} \left[ \mathbf{v}_{nm} \delta_{nm} - \frac{e}{\hbar} (\mathbf{E} \times \Omega_n) \delta_{nm} - \frac{e}{\hbar} \mathbf{E} \cdot \mathbf{r}_{nm,k} \right] a_n^\dagger a_m. \]  \tag{2.33}

The first term is determined by the group velocity of the Bloch electron. The second term is related to anomalous velocity induced by Berry curvature [35]. The modified group velocity is given by

\[ \tilde{\mathbf{v}}_{nn} = \mathbf{v}_{nn} - \frac{e}{\hbar} \mathbf{E} \times \Omega_n. \]  \tag{2.34}

The third term takes into account the modification of band energies with the interband interaction [36], where the effective energy matrix involving bands \( n \) and \( m \) reads

\[ \varepsilon_{nm}(\mathbf{k}, t) = \hbar \omega_n(\mathbf{k}) \delta_{nm} - e \mathbf{r}_{nm}(\mathbf{k}) \cdot \mathbf{E}(t). \]  \tag{2.35}

The effective off-diagonal velocity matrix elements are given by

\[ \tilde{\mathbf{v}}_{nm} = \mathbf{v}_{nm} - \frac{e}{\hbar} \mathbf{r}_{nm,k} \cdot \mathbf{E}. \]  \tag{2.36}

Here the covariant derivative \( \mathbf{r}_{nm,k} = (\mathbf{r}_{nm,x}, \mathbf{r}_{nm,y}, \mathbf{r}_{nm,z}) \) is a vector. We should emphasize that the correction of the off-diagonal velocity itself is not gauge invariant, however, the total product with density matrix \( \rho_{mn} \equiv \left\langle a_n^\dagger a_m \right\rangle / V \) is gauge invariant. We will discuss how to obtain the explicit expression of density matrix by a perturbation theory in the next section.

**2.2.3 Time Dependent Perturbation Theory**

The density operator \( \rho \) obeys the von Neumann equation

\[ i\hbar \frac{d\rho(t)}{dt} = [H(t), \rho(t)] + i\hbar D[\rho(t)], \]  \tag{2.37}
which has a sign difference comparing with Heisenberg equation. \( D[\rho(t)] \) is the relaxation operator which is neglected here in the long time limit. The expectation value of \( A \) is hence given by

\[
\langle A \rangle = \text{tr}(\rho A).
\]

Considering the matrix elements

\[
\langle nk | i\hbar \frac{d\rho}{dt} | mk \rangle = i \hbar \frac{d\rho_{nm}}{dt} = \langle nk | [H, \rho] | mk \rangle = \langle nk | [H_0 - e\mathbf{r} \cdot \mathbf{E}, \rho] | mk \rangle,
\]

where the states with different \( k \) points are assumed to be independent under optical excitation. We then have

\[
\langle nk | [H_0 - e\mathbf{r} \cdot \mathbf{E}, \rho] | km \rangle = \hbar \omega_{nm} \rho_{nm} - e \mathbf{E}(t) \left[ i \rho_{nm,k} + \left( \sum_l (r_{nl} \rho_{lm}) - \sum_l (\rho_{nl} r_{lm}) \right) \right].
\]

The density matrix elements satisfy the relation

\[
\frac{d\rho_{nm}}{dt} = -i \omega_{nm} \rho_{nm} - \frac{e}{\hbar} \mathbf{E}(t) \left[ \rho_{nm,k} - i \left( \sum_l (r_{nl} \rho_{lm}) - \sum_l (\rho_{nl} r_{lm}) \right) \right],
\]

where the diagonal term \( \rho_{nn} \) denotes population and the off-diagonal term \( \rho_{nm} \) with \( n \neq m \) represents coherence. The term \( \rho_{nm,k} \) leads to a coupling between different crystal momenta, which cannot be neglected for nonlinear responses. Scattering effects are absent here. Eq. (2.41) might be considered as collisionless multiband Boltzmann quantum kinetic equation [37, 38]. It should be noted that the density matrix formalism is capable of treating collisional effects which gives rise to the physical line broadening. In general, this equation cannot be solved exactly for physical systems of interest. We therefore seek a practical solution in the form of perturbation expansion. In order to implement this approach, we follow the classic strategy in quantum mechanics and set \( r_{nm} = \lambda r_{nm} \), where \( \lambda \) is the perturbation strength ranging between 0 and 1. The solution of Eq.
(2.41) by means of a power series in parameter $\lambda$ is

$$\rho_{nm} = \rho_{nm}^{(0)} + \lambda \rho_{nm}^{(1)} + \lambda^2 \rho_{nm}^{(2)} + \cdots + \lambda^N \rho_{nm}^{(N)} + \cdots,$$  \hspace{1cm} (2.42)

which recovers the physical scenario when $\lambda = 1$. We thus obtain a set of recursive equations by considering the coefficients of each power parameter $\lambda$ after substituting Eq. (2.42) to the equation Eq. (2.41) and setting $\lambda = 1$

$$\begin{align*}
\frac{d\rho_{nm}^{(0)}}{dt} &= -i\omega_{nm}\rho_{nm}^{(0)} \\
\frac{d\rho_{nm}^{(1)}}{dt} &= -i\omega_{nm}\rho_{nm}^{(1)} - e\hbar^{-1}E(t)\left[\rho_{nm,k} - i\left(\Sigma_l \left(r_{ml}\rho_{lm}^{(0)}\right) - \Sigma_l \left(\rho_{nl}^{(0)} r_{lm}\right)\right)\right] \\
\vdots \\hspace{1cm} \vdots \\
\frac{d\rho_{nm}^{(N)}}{dt} &= -i\omega_{nm}\rho_{nm}^{(N)} - e\hbar^{-1}E(t)\left[\rho_{nm,k}^{(N-1)} - i\left(\Sigma_l \left(r_{ml}\rho_{lm}^{(N-1)}\right) - \Sigma_l \left(\rho_{nl}^{(N-1)} r_{lm}\right)\right)\right]
\end{align*}$$  \hspace{1cm} (2.43)

Assuming the solutions have the following form in the long time limit

$$\rho_{nm}^{(N)}(t) = \sum_{e_1 e_2 \cdots e_N} \rho_{nm}^{(N)e_1 e_2 \cdots e_N} E_{e_1} E_{e_2} \cdots E_{e_N} e^{-i\omega_{\Sigma} t},$$  \hspace{1cm} (2.44)

where $e_N$ and $w_N$ are the indices and frequency of electric field. $\omega_{\Sigma} = \sum_i \omega_{w_i}$ is the sum of all fundamental frequencies. Substituting Eq. (2.44) into Eq. (2.43), we obtain the general solutions for $N$th-order density matrix without the indices of electric field on $\rho_{nm}^{(N)e_1 e_2 \cdots e_N}$

$$\rho_{nm}^{(N)} = \frac{e}{i\hbar(\omega_{nm} - \omega_{\Sigma})} \left[-\rho_{nm}^{(N-1)} + i\sum_l \left(r_{ml} E_{e_l} e_{e_l} - \rho_{nl}^{(N-1)} r_{lm}\right)\right] E_{e_1} E_{e_2} \cdots E_{e_N} e^{-i\omega_{\Sigma} t}. \hspace{1cm} (2.45)$$

We start with the ground state density matrix $\rho_{nm}^{(0)} = f_n \delta_{nm}$. In this case, we recover the well-known Berry phase expression for spontaneous polarization of the unperturbed crystal which is time independent and gauge invariant [33, 39]

$$\langle P(t) \rangle^{(0)} = \langle \Psi | P(t) | \Psi \rangle = \text{tr}(\rho P(t)) = e \int dk \sum_n f_n \xi_{nm}(k), \hspace{1cm} (2.46)$$

16
and obtain the anomalous Hall current [40,41] from the second term in current density Eq. (2.33), which is given by

$$J = -\frac{e^2}{h} E \times \int dk \sum_n f_n \Omega_n(k).$$

(2.47)

This anomalous Hall current vanishes under time-reversal symmetry, and in this case Berry curvature is odd at \( k \) and \( -k \), i.e. \( \Omega(k) = -\Omega(-k) \). The integral of Berry curvature is the Chern number multiplied by a \( 2\pi \) phase factor. Hence, if the time-reversal symmetry is broken, e.g., with external magnetic field, the Hall conductance will take the quantized values \( \sigma = \nu \frac{e^2}{h} \), known as integer quantum Hall effect discovered by Klaus von Klitzing and coauthors [42].

We separate the \( \rho_{nm}^{(N)} \) term into \( \rho_{i,nm}^{(N)} \) and \( \rho_{e,nm}^{(N)} \) originating from \([r_i,\rho]\) and \([r_e,\rho]\), respectively. \( 2^N \) terms are expected when we have \( N \)th order perturbation, which is illustrated by a binary tree diagram as shown in Figure 2.1.

Figure 2.1: A binary tree diagram of perturbative density operators including intraband (blue node) and interband (orange node) terms. At each node, the density operator is separated into the left child (intraband) and the right child (interband).

Substituting the electric field \( E^b(t) = \sum_\beta E_\beta^b e^{-i\omega_\beta t} \) and ground state density matrix into Eq.
\[ \rho^{(1)}_{i,nn} = -\frac{e}{\hbar} \sum_{b\beta} \frac{i}{\omega_\beta} \frac{\partial f_n}{\partial k_b} E^b_\beta e^{-i\omega_\beta t}, \]  

(2.48)

and the interband counterpart

\[ \rho^{(1)}_{e,nm} = \frac{e}{\hbar} \sum_{b\beta} r^b_{nm} f_{mn} \omega_{nm} - \omega_\beta E^b_\beta e^{-i\omega_\beta t}. \]  

(2.49)

Here we have used the following two equations

\[ \langle n k \left[ [r_i, \rho^{(0)}] \right] n k' \rangle = \frac{i}{\hbar} \frac{\partial f_n}{\partial k} \delta(k - k') \]  

(2.50)

\[ \langle n k \left[ [r_e, \rho^{(0)}] \right] m k' \rangle = \left( \sum_l (r_{nl} f_l \delta_{lm}) - \sum_l (\delta_{nl} f_l r_{lm}) \right) \delta(k - k'). \]  

(2.51)

By applying the intraband density matrix \( \rho^{(1)}_{i,nn} \) in the second term in Eq. (2.33), we obtain the Berry curvature induced intraband current

\[ J^{a}_{i,Berry}(-\omega_\Sigma; \omega_\alpha, \omega_\beta, t) = \frac{e^3}{\hbar^2} \int dK \sum_{bcdn,\beta\alpha} \epsilon_{abd} \Omega^d_{n} \frac{i}{\omega_\alpha} \frac{\partial f_n}{\partial k_c} E^b_\beta E^c_\alpha e^{-i\omega_\beta t}, \]  

(2.52)

where \( \omega_\Sigma = \omega_\alpha + \omega_\beta \) is the sum of all the fundamental frequencies. The properties of Levi-Civita symbol \( \epsilon_{abd} \) are provided in APPENDIX A. We again reproduce the semiclassical single band formula for Berry curvature dipole induced anomalous Hall current [28] in the clean limit where \( \tau \to \infty \) and the term \( i/\omega \) can be replaced by \( i/(\omega + i/\tau) = \tau/(1-i\omega\tau) \) with an explicit relaxation time in the semiclassical formula. Eq. (2.52) clearly shows that the intraband second harmonic current is identical to the direct current at zero frequency, say, \( J^{a}_{i,Berry}(-2\omega; \omega, \omega) = J^{a}_{i,Berry}(0; -\omega, \omega) \).

For a linearly polarized light, the current, electric field, and Berry curvature are perpendicular to each other due to the Levi-Civita symbol \( \epsilon_{abd} \). The current is perpendicular to optical field for a circularly polarized light. We will discuss the symmetry consequences for nonlinear anomalous current in Chapter 6. Similarly, we substitute the first order interband density matrix into the
polarization and have

\[
\langle P_{\text{inter}}(t) \rangle^{(1)} = e \int d\mathbf{k} \sum_{nm} r_{nm} \rho_{e,nn}^{(1)} = \frac{e^2}{\hbar} \int d\mathbf{k} \sum_{nm\beta} \frac{f_{nm} r_{nm}^a r_{nm}^b}{(\omega_{mn} - \omega_{\beta} - i\eta)} E_{\beta}^b e^{-i\omega_{\beta}t},
\]

(2.53)

where the imaginary part \(i\eta\) is explicitly added to the frequency. We thus obtain the linear susceptibility. Furthermore, by substituting the first order interband density matrix into the third current term in Eq. (2.33) we obtain the shift current response

\[
J_{e,\text{shift}}^i = -\frac{e^2}{\hbar} \int d\mathbf{k} \sum_{nm} \mathbf{E} \cdot \mathbf{r}_{nm,a} \rho_{e,nn}^{(1)} = -\frac{e^3}{\hbar^2} \int d\mathbf{k} \sum_{bcmn\alpha\beta} \frac{r_{nm,a}^c r_{mn}^f f_{nm}}{\omega_{mn} - \omega_{\beta}} E_{\beta}^b E_{\alpha}^c e^{-i\omega_{\beta}t}.
\]

(2.54)

Under the second order perturbation, the polarization and current responses are more sophisticated as they involve four density matrix terms. We substitute the first order intraband density matrix \(\rho_{i,nn}^{(1)}\) into Eq. (2.43) and get the second order intraband density matrix

\[
\rho_{ii,nn}^{(2)} = -\frac{e}{\hbar^2} \sum_{a} \frac{i}{\omega_{\alpha}} \frac{\partial \rho_{i,nn}^{(1)}}{\partial \mathbf{k}_a} E_{\alpha}^a e^{-i\omega_{\alpha}t} = -\frac{e^2}{\hbar^2} \sum_{ab\alpha\beta} \frac{1}{\omega_{\alpha}\omega_{\beta}} \frac{\partial^2 f_n}{\partial k_a \partial k_b} E_{\alpha}^a E_{\beta}^b E_{\gamma}^e e^{-i\omega_{\gamma}t}.
\]

(2.55)

This term is an even function at \(\mathbf{k}\) and \(-\mathbf{k}\), therefore has no contribution to the current in time reversal invariant system (TRIS). However, it can be nonvanishing in non-TRIS

\[
J_{i,Berry}^e(-\omega_{\Sigma}; \omega_{\alpha}, \omega_{\beta}, \omega_{\gamma}, t) = \frac{e^4}{\hbar^3} \int d\mathbf{k} \sum_{abcd\alpha\beta} \epsilon_{e\alpha\beta} \Omega_{\gamma}^d \frac{1}{\omega_{\alpha}\omega_{\beta}} \frac{\partial^2 f_n}{\partial k_a \partial k_b} E_{\alpha}^a E_{\beta}^b E_{\gamma}^e e^{-i\omega_{\gamma}t}.
\]

(2.56)

A new direct current response is predicted under static field, say, \(\omega_{\gamma} = 0\), without time reversal symmetry. This current is related to the second derivative of Berry curvature, referred to as Berry curvature quadruple. Using the same techniques, we can obtain the density matrix

\[
\rho_{ie,nn}^{(2)} = \frac{e}{\hbar \omega_{\Sigma}} \sum_{l\alpha\alpha} E_{\alpha}^a \left( r_{nl}^a \rho_{e,ln}^{(1)} - \rho_{e,ln}^{(1)} r_{nl}^a \right)
= \frac{e^2}{\hbar^2 \omega_{\Sigma}} \sum_{ab\alpha\beta} r_{nl}^a r_{ln}^b f_{nl} \left( \frac{1}{\omega_{ln} - \omega_{\beta}} + \frac{1}{\omega_{nl} - \omega_{\beta}} \right) E_{\alpha}^a E_{\beta}^b e^{-i\omega_{\beta}t}.
\]

(2.57)
Substituting this density matrix in Eq. (2.33) for the first current term, we obtain the injection current
\[
\frac{dJ_{a}^{\text{inj}}}{dt} = \frac{-ie^{3}}{\hbar^{2}} \int d\mathbf{k} \sum_{bemn} f_{mn} \Delta_{mn}^{a} \gamma_{nm}^{b} \gamma_{mn}^{c} \frac{1}{\omega_{nm} - \omega_{\beta}} E_{\alpha}^{b} E_{\gamma}^{c} e^{-i\omega_{\Sigma} t} = \frac{\pi e^{3}}{\hbar^{2}} \int d\mathbf{k} \sum_{bemn} f_{mn} \Delta_{mn}^{a} \gamma_{nm}^{b} \gamma_{mn}^{c} \delta(\omega_{nm} - \omega_{\beta}) E_{\alpha}^{b} E_{\gamma}^{c} e^{-i\omega_{\Sigma} t}.
\]
(2.58)
We will discuss more details in Chapter 4 and Chapter 6. Now we solve the second order density matrix based on the first order interband term, and we have
\[
\rho_{ee, mn}^{(2)} = \frac{e}{i\hbar(\omega_{mn} - \omega_{\Sigma})} \left[ -\rho_{e, mn, \alpha}^{(1)b} + i \sum_{\pi, \pi, \pi} \left( \rho_{m, \alpha}^{(1)b} \rho_{n, \pi}^{(1)b} - \rho_{e, \pi, \pi}^{(1)b} \rho_{e, \pi, \pi}^{(1)b} \right) \right] E_{\beta}^{b} E_{\gamma}^{c} e^{-i\omega_{\Sigma} t},
\]
(2.59)
and
\[
\rho_{ei, mn}^{(2)} = \frac{e}{\hbar(\omega_{mn} - \omega_{\Sigma})} (\rho_{m, \pi}^{(1)b} - \rho_{e, \pi, \pi}^{(1)b}) E_{\beta}^{b} E_{\gamma}^{c} e^{-i\omega_{\Sigma} t}.
\]
(2.60)
Let’s simplify the mix term \(\rho_{ei, mn}^{(2)}\)
\[
\rho_{ei, mn}^{(2)} = -\frac{e^{2} r_{mn}^{e}}{\hbar^{2} (\omega_{mn} - \omega_{\Sigma})} \rho_{m, \pi}^{(1)b} - \rho_{e, \pi, \pi}^{(1)b} \frac{i}{\omega_{\beta}} \frac{\partial f_{mn}}{\partial k_{b}} E_{\beta}^{b} E_{\gamma}^{c} e^{-i\omega_{\Sigma} t},
\]
(2.61)
and substitute \(\rho_{ei, mn}^{(2)}\) into interband current
\[
J^{a} = \frac{e^{3}}{\hbar^{2}} \int d\mathbf{k} \frac{r_{mn}^{a} r_{mn}^{e}}{(\omega_{mn} - \omega_{\Sigma})} \frac{\omega_{nm} \partial f_{mn}}{\partial k_{b}} E_{\beta}^{b} E_{\gamma}^{c} e^{-i\omega_{\Sigma} t}.
\]
(2.62)
In the limit of direct current response \(\omega_{\Sigma} = 0\), we arrive at
\[
J^{a} = \frac{e^{3}}{\hbar^{2}} \int d\mathbf{k} \sum_{bemn} r_{mn}^{a} r_{mn}^{e} \frac{\partial f_{mn}}{\partial k_{b}} \frac{1}{\omega_{\beta}} E_{\beta}^{b} E_{\gamma}^{c},
\]
(2.63)
which is similar to intraband Berry curvature dipole current. The optical rectification, second harmonic generation, and higher order NLO effects can also be derived from this general density matrix framework. Other theories will be discussed in following sections.
2.3 Floquet Theory, Green’s function and TDDFT

The dynamics of quantum systems under intense time-dependent external fields are explicitly described by time-dependent Hamiltonians. Such systems exhibit an interplay of nonlinearity, nonequilibrium behavior, and quantum tunneling [43]. We start with the Hamiltonian with a discrete time translation symmetry

\[ H(t) = H(t + T), \]

(2.64)

where \( T \) is the period of the time-dependent perturbation. The Schrödinger equation is given by

\[ \left( H(\mathbf{r}, t) - i\hbar \frac{\partial}{\partial t} \right) \Psi(\mathbf{r}, t) = 0. \]

(2.65)

The solution of the periodic time-dependent Schrödinger equation is equivalent to the Bloch theorem [44,45] (so-called Floquet formalism)

\[ \Psi(\mathbf{r}, t) = e^{-i\varepsilon t} \phi(\mathbf{r}, t), \phi(\mathbf{r}, t) = \phi(\mathbf{r}, t + T). \]

(2.66)

Here, \( \varepsilon \) is the quasienergy, being equal to multiples of \( \hbar \omega = \hbar 2\pi / T \), e.g., the energy of incident photons. One finds

\[ (H(t) - i\hbar \partial_t) \phi(\mathbf{r}, t) = \varepsilon \phi(\mathbf{r}, t). \]

(2.67)

For periodic functions of period, \( T \), the Fourier series expansion may be written by

\[ f(t) = \sum_{q=-\infty}^{\infty} F_q e^{-iqt}, \]

(2.68)

where the Fourier coefficient has the same unit as \( f(t) \), given by the Fourier integral

\[ F_q = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{iqwt} dt = \frac{1}{T} \int_{0}^{T} f(t) e^{iqwt} dt. \]

(2.69)
Using the Fourier transformation, we obtain

\[ \phi(t) = \sum_m e^{-im\omega t} \phi_m. \]  \hspace{1cm} (2.70)

We then have

\[ (i\hbar \partial_t + \varepsilon) \phi(t) = \sum_m (m\hbar \omega + \varepsilon) e^{-im\omega t} \phi_m = H(t) \phi(t) = \sum_q H_q e^{-i\omega t} \sum_n e^{-in\omega t} \phi_n \]

\[ = \sum_q \sum_n H_q e^{-i(q+n)\omega t} \phi_n \xrightarrow{q \to m-n} \sum_{m-n} \left( \sum_n H_{mn} \phi_n \right) e^{-im\omega t} \]  \hspace{1cm} (2.71)

\[ = \sum_m \left( \sum_n H_{mn} \phi_n \right) e^{-im\omega t}, \]

where

\[ H_{mn} = \frac{1}{T} \int_0^T H(t) e^{i(m-n)\omega t} dt. \]  \hspace{1cm} (2.72)

With the above matrix elements, it’s easy to show that

\[ (m\hbar \omega + \varepsilon) \phi_m = \sum_n H_{mn} \phi_n. \]  \hspace{1cm} (2.73)

Finally, we obtain the Floquet formalism in a matrix form of time-dependent Schrödinger equation

\[ H_F \phi = \varepsilon \phi, \]  \hspace{1cm} (2.74)

where

\[ (H_F)_{mn} = H_{mn} - m\hbar \omega \delta_{mn}. \]  \hspace{1cm} (2.75)

The eigenvalues of Floquet Hamiltonian give the Floquet side bands of dressed states. The translation symmetry of time is effectively incorporated to the Floquet index of quasienergy. The time-dependent current in terms of velocity operator matrix element in the steady Floquet state is given by

\[ J^a(t) = \sum_{mn} \left\{ \text{tr} \left[ \langle \psi | \hat{v}^a | \psi \rangle \right] \right\}_{mn} e^{-i(m-n)\omega t}. \]  \hspace{1cm} (2.76)
The Hamiltonian can be expanded to the power series of vector potential

\[ H(\vec{p} - eA(t)) = H_0 - \sum_i (\partial_{k_i} H_0) eA_i e^{-i\omega t} + \sum_{i,j} \frac{1}{2} (\partial_{k_i} \partial_{k_j} H_0) e^2 A_i A_j e^{-2i\omega t} + \ldots \]  

(2.77)

Assuming the perturbed Hamiltonian has a general form

\[ H(t) = H_0 + V e^{i\Omega t} + V^\dagger e^{-i\Omega t}, \]  

(2.78)

we can obtain the time independent Floquet Hamiltonian

\[
H_F = \begin{pmatrix}
\ddots & \vdots & \vdots & \vdots & \ddots \\
\vdots & H_0 + 2\hbar \Omega I & V & 0 & 0 \\
\vdots & V^\dagger & H_0 + \hbar \Omega I & V & 0 \\
\vdots & 0 & V^\dagger & H_0 & V \\
\ddots & 0 & 0 & V^\dagger & H_0 - \hbar \Omega I \\
\ddots & \vdots & \vdots & \vdots & \ddots 
\end{pmatrix}.
\]  

(2.79)

Oka and Aoki proposed a TKNN-type formula for the Hall conductivity based on the Floquet states [46], where a circularly polarized light opens a gap in the Dirac cone in graphene. The Hall conductivity is given by

\[
\sigma_{xy}(A) = -\frac{e^2}{\hbar} \int d\vec{k} \sum_\alpha f_\alpha(k) [\nabla_k \times A_\alpha(k)]_z,
\]  

(2.80)

where \( \alpha \) includes original band and Floquet index. Morimoto and coauthors also studied nonlinear magneto-optical responses of metals and reproduced the Berry curvature dipole induced nonlinear current in the many-band case [47]. The Floquet formalism is a very useful quantum mechanical tool to investigate NLO phenomena. In addition, periodically driven quantum system can modify the long-time dynamics in a profound manner, activate novel topological order, and enable novel physics that are inaccessible in static systems [48].
Green’s function describes the probability of finding an electron with spin $\sigma$ at position $r$ and time $t$. At zero temperature, the one-particle Green’s function is defined as

$$G(\mathbf{r}_t, \mathbf{r}'_{t'}) = -\frac{i}{\hbar} \langle 0 | \hat{T} \psi(\mathbf{r}_t) \psi^\dagger(\mathbf{r}'_{t'}) | 0 \rangle,$$

(2.81)

where $\hat{T}$ is so called Wick’s time ordering operator. The equation of motion for Green’s function is given by

$$\left[ i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m_e} \nabla^2 - \varphi(\mathbf{r}_t) \right] G(\mathbf{r}_t, \mathbf{r}'_{t'}) = \delta(\mathbf{r}_t, \mathbf{r}'_{t'})$$

$$-\frac{i}{\hbar} \int d\mathbf{r}'' d\mathbf{r}''' v(\mathbf{r}_t, \mathbf{r}''_{t''}) \langle \hat{T} n(\mathbf{r}''_{t''}) \psi(\mathbf{r}_t) \psi^\dagger(\mathbf{r}'_{t'}) \rangle,$$

(2.82)

where $\hat{n}$ is the charge density operator and $v$ is the Coulomb interaction operator. Here $v(\mathbf{r}_t, \mathbf{r}''_{t''}) = v(\mathbf{r}, \mathbf{r}'') \delta(t - t'')$. $\varphi$ is the scalar electric potential including the effect from both electronic and nuclear charges. One may use the functional derivative \cite{49} to rewrite the equation of motion \cite{50}

$$\left[ i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m_e} \nabla^2 - V_{\text{tot}}(\mathbf{r}_t) - i\hbar \int d\mathbf{r}'' d\mathbf{r}''' v(\mathbf{r}_t + \eta, \mathbf{r}''_{t''}) \times \frac{\delta}{\delta \varphi(\mathbf{r}''_{t''})} \right] G(\mathbf{r}_t, \mathbf{r}'_{t'}) = \delta(\mathbf{r}_t, \mathbf{r}'_{t'}),$$

(2.83)

where

$$V_{\text{tot}}(\mathbf{r}_t) = \int d\mathbf{r}' v(\mathbf{r}, \mathbf{r}') \langle \hat{n}(\mathbf{r}'_t) \rangle + \varphi(\mathbf{r}_t).$$

(2.84)

We use the common convention of space-spin-time, where spin labels are implied, that is, $(\mathbf{r}_t) \rightarrow 1$, $(\mathbf{r}'_{t'}) \rightarrow 2$, and $(\mathbf{r}_t + \eta) \rightarrow 1^+$. The equation is given by

$$\left[ i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m_e} \nabla^2 (1) - V_{\text{tot}}(1) - i\hbar \int d(3) v(1^+ 3) \frac{\delta}{\delta \varphi(3)} \right] G(12) = \delta(12).$$

(2.85)

We have Kadanoff-Baym equation

$$\frac{\delta G(12)}{\delta \varphi(3)} = -\int d(45) G(14) \frac{\delta G^{-1}(45)}{\delta \varphi(3)} G(52),$$

(2.86)

and the chain rule

$$\frac{\delta G^{-1}(45)}{\delta \varphi(3)} = \int d(6) \frac{\delta G^{-1}(45)}{\delta V_{\text{tot}}(6)} \frac{\delta V_{\text{tot}}(6)}{\delta \varphi(3)}.$$

(2.87)
One obtains a more compact form of the equation

\[
\left[ i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m_e} \nabla^2(1) - V_{\text{tot}}(1) \right] G(12) - \int d(3) \Sigma(13) G(32) = \delta(12), \tag{2.88}
\]

where

\[
\Sigma(12) = i\hbar \int d(34) G(13) \Gamma(324) W(41^+) \tag{2.89}
\]

is the electron self-energy including many-body effects due to exchange and correlation. Here \( \Gamma(123) \) is vertex function, defined by

\[
\Gamma(123) \equiv -\delta G^{-1}(12)/\delta V_{\text{tot}}(3), \tag{2.90}
\]

and

\[
W(12) \equiv \int d(3) \epsilon^{-1}(13) v(32) = \int d(3) v(13) \epsilon^{-1}(23) \tag{2.91}
\]

is the screened Coulomb interaction. \( \epsilon^{-1} \) is the many-body dielectric function given by

\[
\epsilon^{-1}(12) = \delta V_{\text{tot}}(1)/\delta \varphi(2). \tag{2.92}
\]

The relation between the non-interacting Green’s function \( G_0 \) and the interacting Green’s function \( G \) is given by the Dyson equation

\[
G^{-1}(\mathbf{r}, \mathbf{r}', \omega) = G_0^{-1}(\mathbf{r}, \mathbf{r}', \omega) - \Sigma(\mathbf{r}, \mathbf{r}', \omega). \tag{2.93}
\]

Using non-interacting Kohn-Sham (KS) Green’s function \( G_{KS} \) for \( G_0 \), we have

\[
G(12) = G_{KS}(12) + \int d(34) G_{KS}(13) \Sigma(34) G(42). \tag{2.94}
\]

In fact, Hedin-Baym equations give a rigorous Green’s function formalism for investigating interacting electrons and phonons in materials from first-principles. For interacting electrons, we
rewrite the Hedin’s equations \[51\]

\[
P(12) = -i \int d(34) G(13) G(41^+) \Gamma(342),
\]

\[
W(12) = \nu(12) + \int d(34) W(13) P(34) \nu(42),
\]

\[
\Sigma(12) = i \int d(34) G(14^+) W(13) \Gamma(423),
\]

\[
\Gamma(123) = \delta(12) \delta(13) + \int d(4567) \frac{\delta \Sigma(12)}{\delta G(45)} G(46) G(75) \Gamma(673).
\]

In principle, the Hedin’s equations can be solved self-consistently. However, it’s too difficult to solve in a non-perturbative manner. The GW approximation (named after the form of the self-energy operator) is commonly used, where the vertex function is approximated by

\[
\Gamma(123) = \delta(12) \delta(13).
\]

Hence the self-energy becomes

\[
\Sigma(12) = iG(12^+) W(12).
\]

In this approximation, the irreducible polarizability is given by

\[
P(12) = -iG(12) G(21^+),
\]

which is equal to random phase approximation without considering electron-hole interactions. To go beyond the GW approximation, two-particle Bethe-Salpeter equation (BSE) can be utilized which includes the electron-hole interaction (excitonic effect). Now using the approximation \(\delta \Sigma / \delta G = iW\) which neglects the excitation induced change in the screening, one obtains

\[
\Gamma(123) = \delta(12) \delta(13) iW (1^+2) \int d(67) G(16) G(72) \Gamma(673).
\]
Hence one can define a generalized three point polarization

\[ 3P(312) \equiv -i \int d(67)G(16)G(72)\Gamma(673). \]  

(2.100)

Similarly, one can further define the four-point screened Coulomb interaction

\[ 4W(1234) \equiv W(12)\delta(13)\delta(24), \]  

(2.101)

and obtain the BSE for four-point polarization

\[ 4\overline{P} = 4P_{IQP} + 4P_{IQP}K4\overline{P}, \]  

(2.102)

where the kernel function \( K \) includes an electron-hole exchange \( v \) and electron-hole attraction \( W \)

\[ K(1234) = \delta(12)\delta(34)\overline{v}(13) - \delta(13)\delta(24)W(12). \]  

(2.103)

BSE based exchange-correlation kernel has also been introduced for periodic system.

It should be noted that the time-dependent density-functional theory (TDDFT) might be less cumbersome compared with the BSE formalism in the case that TDDFT is based on electron density involving two-point response functions while BSE relies on Green’s function with four-point responses. Time-dependent KS equation [52] was first derived by Runge and Gross. One can define noninteracting electronic system that obey the time-dependent KS equation

\[ i\frac{\partial \phi_j(r,t)}{\partial t} = \left[ -\frac{\nabla^2}{2} + v_{KS}[n](r,t) \right] \phi_j(r,t), \]  

(2.104)

where the density is given by

\[ n(r,t) = \sum_{i=1}^{N} |\phi_j(r,t)|^2. \]  

(2.105)

Similar to the KS equation, the potential can be separated into ionic, Hartree, and exchange-
correlation part
\[ v_{KS}(\mathbf{r}, t) = v_{\text{ext}}(\mathbf{r}, t) + v_{H}(\mathbf{r}, t) + v_{xc}(\mathbf{r}, t). \] (2.106)

The adiabatic local density approximation (ALDA) is the simplest approximation for the time-dependent exchange-correlation functional, that is
\[ v^{\text{ALDA}}_{xc}[n](\mathbf{r}, t) = v^{\text{hom}}_{xc}(n(\mathbf{r}, t)) = \frac{d}{dn} \left[ n e^{\text{hom}}_{xc}(n) \right] \bigg|_{n=n(\mathbf{r}, t)}, \] (2.107)

where \( e^{\text{hom}}_{xc}(n) \) is the exchange-correlation energy per particle of the homogeneous electron gas.

Solving time-dependent KS equation, one can directly determine optical properties from time-dependent total dipole moment
\[ \mathbf{p}(t) = \int \rho(\mathbf{r}, t) \mathbf{r} d^3 \mathbf{r}. \] (2.108)

The time-dependent polarization can be expanded in power series of electric field
\[ p_i(t) = p_0 + \int dt_1 \chi^{(1)}_{ij}(t; t_1) E_j(t_1) + \int dt_1 \int dt_2 \chi^{(2)}_{ijk}(t; t_1, t_2) E_j(t_1) E_k(t_2) + \int dt_1 \int dt_2 \int dt_3 \chi^{(3)}_{ijkl}(t; t_1, t_2, t_3) E_j(t_1) E_k(t_2) E_l(t_3) + \cdots \] (2.109)
\[ \equiv p_i^{(0)} + p_i^{(1)}(t) + p_i^{(2)}(t) + p_i^{(3)}(t) + \cdots \]

It should be emphasized that the ALDA functional works reasonably well for metals, but not very good for semiconductors and insulators. Nevertheless, TDDFT is still promising as it can go beyond the perturbative regime which cannot be easily treated in the GW approximation under strong external fields. For extended periodic systems in a uniform electric field, time-dependent current density functional theory (TDCDFT) has been proposed which includes vector potential in Hamiltonian
\[ H = \sum_i \left\{ \frac{1}{2m} \left[ (\mathbf{p}_i + \frac{e}{c} \mathbf{A}_i)^2 \right] + v_{\text{ext},i} \right\} + v_{ee}. \] (2.110)

The vector potential is uniquely determined by the time-dependent current density under a gauge
transformation [53]. This approach can be applied to the laser induced NLO responses in periodic systems.

2.4 Tight-Binding Method and Wannier Representation

The tight-binding method is a practical and efficient approach to calculate electronic band structure and geometric properties of solids in a localized basis set. Wannier functions (WFs) are one well-known example which are localized in real space. WFs can be numerically constructed from the Bloch wavefunctions, which enables efficient first-principles calculations for many physical quantities. Following the convention and method by Wang and coauthors [54], we write the Bloch-like basis function in the Wannier gauge in terms of the phased sum of WFs

$$| u_{nk}^{(W)} \rangle = \sum_{R} e^{-i\mathbf{k} \cdot (\mathbf{r} - \mathbf{R} - \mathbf{\tau}_n)} | \mathbf{R} n \rangle,$$

(2.111)

where \( \mathbf{R} \) denotes the unit cell and \( n \) labels the WFs, and \( \mathbf{\tau}_n = \langle 0n | \mathbf{r} | 0n \rangle \). The Bloch eigenfunction in the Hamiltonian gauge can be expended as a linear combination of Bloch-like basis in the Wannier gauge

$$| u_{nk} \rangle = | u_{nk}^{(H)} \rangle = \sum_{m} | u_{mk}^{(W)} \rangle U_{mn}(k).$$

(2.112)

Electronic and topological quantities can be expressed analytically in the Wannier gauge as the following

$$H_{nm,\alpha}^{(W)}(k) = \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} \langle 0n | \hat{H} | \mathbf{R} m \rangle$$

$$A_{nm,\alpha}^{(W)}(k) = \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} i R_{\alpha} \langle 0n | \hat{r}_\alpha | \mathbf{R} m \rangle$$

$$\Omega_{nm,\alpha\beta}^{(W)}(k) = \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} (i R_{\alpha} \langle 0n | \hat{r}_\beta \rangle | \mathbf{R} m \rangle - i R_{\beta} \langle 0n | \hat{r}_\alpha \rangle | \mathbf{R} m \rangle).$$

(2.113)
A gauge transformation of the quantities, e.g., dipole matrix element is applied, which leads to

\[
A^{(H)}_\alpha = i \left( u^{(H)}_{nk} | \partial_\alpha | u^{(H)}_{nk} \right) = i \left( \sum_m | u^{(W)}_{mk} \rangle U_{mn}(k) \right)^\dagger \partial_\alpha \left( \sum_{m'} | u^{(W)}_{m'k} \rangle U_{m'n}(k) \right)
= U^\dagger A^{(W)}_\alpha U + iU^\dagger \partial_\alpha U.
\] (2.114)

The derivative \( \partial_\alpha U \) can be obtained from perturbation theory

\[
\partial_\alpha U_{mn} = \sum_l U_{ml} D^{(H)}_{lm,\alpha} = \left( U D^{(H)}_{\alpha} \right)_{mn},
\] (2.115)

where

\[
D^{(H)}_{nm,\alpha} = \begin{cases} 
\frac{P^{(H)}_{nm,\alpha}}{\epsilon^{(H)}_m - \epsilon^{(H)}_n} & \text{if } n \neq m \\
0 & \text{if } n = m,
\end{cases}
\] (2.116)

and \( R^{(H)}_{nm,\alpha} = \left( U^\dagger H^{(W)}_{\alpha} U \right)_{nm} \). We then obtain Berry connection in the Hamiltonian gauge

\[
A^{(H)}_\alpha = \tilde{A}^{(H)}_\alpha + iD^{(H)}_\alpha,
\] (2.117)

and Berry curvature

\[
\Omega^{(H)}_{\alpha\beta} = \tilde{\Omega}^{(H)}_{\alpha\beta} - \left[ D^{(H)}_{\alpha\beta}, \tilde{A}^{(H)}_\beta \right] + \left[ D^{(H)}_{\beta\alpha}, \tilde{A}^{(H)}_\alpha \right] - i \left[ D^{(H)}_{\alpha\beta}, D^{(H)}_\beta \right].
\] (2.118)

The covariant derivative can be calculated using the sum rule

\[
\hat{r}^a_{nm;b} = \frac{i}{\omega_{nm}} \left[ \frac{v^a_{nm} \Delta^b_{nm} + v^b_{nm} \Delta^a_{nm}}{\omega_{nm}} + \sum_{p+n,m} \frac{v^a_{np} v^b_{pm} - v^b_{np} v^a_{pm}}{\omega_{pm}} - \omega_{nm} \right] (m \neq n),
\] (2.119)

where \( w^{ab}_{nm} = \frac{1}{n} \left( u_n | \partial^2 H | u_m \right) \) is the mass term, which includes the nonlocal effect properly in first-principles based tight-binding Hamiltonian. The derivation of the sum rule is shown in APPENDIX
A. For local potential $V = V(r)$, the mass term is given by

$$
\partial_a \partial_b H(k) = \partial_a \partial_b \left( e^{-ikr} H(r) e^{ikr} \right) = \partial_a \left( -ie^{-ikr} \left[ r^b, H \right] e^{ikr} \right) = \partial_a \left( \hbar e^{-ikr} v^b e^{ikr} \right) = \partial_a \left( -ie^{-ikr} \left[ r^a, v^b \right] e^{ikr} \right) = \frac{\hbar^2}{m} \delta_{ab}.
$$

(2.120)

The tight-binding approach enables much more efficient calculation of NLO responses in topological materials where a highly dense k-point sampling of the first Brillouin zone is often required.

2.5 Group Theory

Franz Neumann’s principle states that physical properties must be invariant under the symmetry operations $\hat{P}_R = (R|\tau + a)$ in the space (or point) group of crystal. Symmetry makes great simplifications even without giving any details of the Hamiltonian. As the symmetry operator $\hat{P}_R$ commutes with Hamiltonian $H$, we have

$$
\hat{P}_R H \psi_n = \hat{P}_R E_n \psi_n = H \left( \hat{P}_R \psi_n \right) = E_n \left( \hat{P}_R \psi_n \right).
$$

(2.121)

Thus $\hat{P}_R \psi_n$ is also an eigenfunction of Hamiltonian $H$. Specifically, if we apply the translation operation $\hat{P}_{\{\varepsilon|\tau\}}$ on wavefunction $\psi_k(r)$

$$
\hat{P}_{\{\varepsilon|\tau\}} \psi_k(r) = \psi_k(r + \tau) = e^{ik\cdot\tau} \psi_k(r),
$$

(2.122)

where $\tau$ can be written in terms of the sum of unit vectors in real space

$$
\tau = \sum_{i=1}^{3} n_i a_i.
$$

(2.123)

We can recover the Bloch’s theorem from the fact that the translation group is Abelian [55]

$$
\psi_k(r) = e^{ik\cdot\tau} u_k(r).
$$

(2.124)
Another important symmetry operation is time reversal, transforming the wavefunction to its complex one with opposite $k$ and spin

$$\hat{T}\psi_{n,k\uparrow}(r) = \psi_{n,-k\downarrow}(r).$$

(2.125)

Hence, the time reversal symmetry gives $E_{n\downarrow}(k) = E_{n\uparrow}(-k)$, and inversion symmetry gives $E_n(k) = E_n(-k)$. This is the well-known Kramers degeneracy.

In solid state physics, physical properties of solids can be generally regarded as tensors [56]. The permutation symmetry classifies the tensors in an effective way. For example, in linear susceptibility, we have the permutation relation $\chi_{ij} = \chi_{ji}$, which reduces 9 independent components into 6. The number of independent non-zero tensor components is determined the direct product of the irreducible representation $\Gamma_i$ of fields and responses

$$\{\Gamma_J\} \otimes \{\Gamma_{E\ldots}\} = \sum_i \alpha_i \Gamma_i \otimes \sum_j \beta_j \Gamma_j = \sum_k \gamma_k \Gamma_k.$$

(2.126)

Let’s take the hexagonal point group symmetry $D_{6h}$ as an example. The linear susceptibility tensor requires the direct product

$$\Gamma_P \otimes \Gamma_E = (A_{2u} + E_{1u}) \otimes (A_{2u} + E_{1u})$$

$$= 2A_{1g} + A_{2g} + 2E_{1g} + E_{2g}$$

(2.127)

$$= (2A_{1g} + E_{1g} + E_{2g})^{(s)} + (A_{2g} + E_{1g})^{(a)},$$

implying that there are only two independent components for linear susceptibility tensor in $D_{6h}$ systems since there are two total symmetric representations $A_{1g}$ in the product. For the second order nonlinear optical susceptibility tensor, we have

$$\Gamma_P \otimes \Gamma_{EE}^{(s)} = (A_{2u} + E_{1u}) \otimes (2A_{1g} + E_{1g} + E_{2g})$$

$$= (A_{1u} + A_{2u} + B_{1u} + B_{2u} + 2E_{1u} + E_{2u})^{(s)} + (2A_{2u} + 4E_{1u} + E_{2u})^{(a)},$$

(2.128)
indicating there are no non-vanishing terms, that is, all second order nonlinear susceptibilities (related to third rank tensor) vanish in $D_{6h}$ systems. This is consistent with inversion symmetry present in $D_{6h}$ point group where all second order optical nonlinearities are absent. Just as in the words of Steven Weinberg, “the universe is an enormous direct product of representations of symmetry groups.” Here, group theory is a powerful tool to investigate NLO phenomena in a qualitative manner.
3. DISCOVERY OF 2D MULTIFERROICS AND TOPOLOGICAL MATERIALS

3.1 Two-Dimensional Multiferroics in Monolayer Group IV Monochalcogenides *

Low-dimensional multiferroic materials hold great promises in miniaturized device applications such as nanoscale transducers, actuators, sensors, photovoltaics, and nonvolatile memories. Here, using first-principles theory we predict that two-dimensional (2D) monolayer Group IV monochalcogenides including GeS, GeSe, SnS, and SnSe are a class of 2D semiconducting multiferroics with strongly coupled giant in-plane spontaneous ferroelectric polarization and spontaneous ferroelastic lattice strain that are thermodynamically stable at room temperature and beyond, and can be effectively modulated by elastic strain engineering. Their optical absorption spectra exhibit strong in-plane anisotropy with visible-spectrum excitonic gaps and sizable exciton binding energies, rendering the unique characteristics of low-dimensional semiconductors. More importantly, the predicted low domain wall energy and small migration barrier together with the coupled multiferroic order and anisotropic electronic structures suggest their great potentials for tunable multiferroic functional devices by manipulating external electrical, mechanical, and optical field to control the internal responses, and enable the development of four device concepts including 2D ferroelectric memory, 2D ferroelastic memory, and 2D ferroelastoelectric nonvolatile photonic memory as well as 2D ferroelectric excitonic photovoltaics.

3.1.1 Brief Introduction to 2D Materials and Multiferroics

The last decade has witnessed tremendous progress in two-dimensional (2D) materials [57–62] research, leading to the remarkable discoveries of 2D metals, semiconductors, and insulators and, equally important, a variety of rich physics such as massless [59] and massive Dirac fermions [63] and topological insulators [64, 65]. Recently, coupled physical properties of 2D materials attract a lot of attention for their potentials device and energy applications. For example, monolayer BN [66], MoS$_2$ [67–69] and Group IV monochalcogenides [70, 71] were recently found to pos-

sess large piezoelectricity [72] which allows efficient mechanical-to-electrical energy conversion. Monolayer CrSiTe$_3$ was experimentally synthesized and demonstrated with 2D ferromagnetic ordering [73], paving the way towards 2D spintronics. Interestingly, Peierls-distorted 1T’ transition metal chalcogenide have been predicted to exhibit both quantum spin Hall effect with a nontrivial $Z_2$ topological index [74,75] and ferroelasticity [76], suggesting the possibility of controlling the anisotropy of the topologically-protected edge states via elastic strain engineering [77].

2D materials mentioned above possess one ferroic order only. Undoubtedly, 2D multiferroic materials that hold simultaneously two or more primary ferroic (i.e., ferroelectric, ferromagnetic, and ferroelastics) orders are highly desirable as they will open up unprecedented opportunities for both scientific and technological endeavors [78–84]. Ideal multiferroics, however, require different order parameters to be strongly coupled, which in tandem with appropriate kinetic barrier of phase transition will allow facile switching of electric polarization, magnetization, and lattice strain through external disparate fields. Though highly valuable, perfect multiferroic materials especially in low dimensions are scarce, largely due to the stringent symmetry and chemistry requirements for the coexistence of two or more coupled ferroic orders and good thermodynamic stability for practical applications at room temperature [78,82].

Based on first-principles theory, here we predict that monolayer Group IV monochalcogenides (including GeS, GeSe, SnS, and SnSe) represent a class of 2D multiferroic materials that simultaneously possess strongly-coupled ferroelectric and ferroelastic orders and, more importantly, have low domain wall energy and small migration barrier. The calculated giant intrinsic in-plane electric polarization implies that their thickness is not constrained by out-of-plane depolarization-induced instability that often occurs in free-standing multiferroic ultrathin films [85]. As ferroelastic order in these monolayers is pertinent to spontaneous strain of unit cell along two perpendicular orientations, the direct coupling between ferroelastic lattice strain and ferroelectric polarization allows the direct control of one ferroic order by applying external field that is conjugated with the other one. Furthermore, the spontaneous lattice strain and in-plane polarization also lead to highly anisotropic electronic and optical properties. Therefore, these 2D ferroelastoelectric semicon-
ducting monolayers with highly anisotropic, strongly coupled, and externally switchable physical properties will engender a wide variety of ultrathin mechano-opto-electronic applications. Applications of elastic strain, electric field, or optical field can efficiently switch the multiferroic states and alter their electric, optical, and mechanical responses, thereby enabling the conceptual designs of 2D ferroelectric, ferroelastic, and ferroelastoelectric nonvolatile photonic memories as well as 2D ferroelectric excitonic photovoltaics presented below.

3.1.2 Ferroelastic Order and Spontaneous Strain in Monolayer MX

Figure 3.1: Structure of monolayer Group IV monochalcogenides (MX) and their ferroelastic and ferroelectric orders. (a) Perspective view of free-standing MX monolayer using GeSe as an example. (b) and (c): two ferroelectric states with opposite spontaneous polarization along the $x$ axis: $-P_x$ and $+P_x$. (d) and (e): two ferroelectric states with opposite spontaneous polarization along the $y$ axis: $-P_y$ and $+P_y$. (b,c) and (d,e): two ferroelastic states with spontaneous tensile strain along $x$ and $y$ relative to the centrosymmetric paraelastic structure. The corresponding transformation strain matrices are denoted by $\eta_x$ and $\eta_y$. 
Monolayer Group IV monochalcogenide (abbreviated as MX), as shown in Figure 3.1a, consists of two puckered atomic layers similar to monolayer black phosphorus where M=(Ge, Sn) and X=(S, Se). We use GeSe as an example and set the $x$-axis and $y$-axis as in-plane axes with the $z$-axis along the plane normal. Its noncentrosymmetric unit cell is illustrated by dashed orange rectangle in Figures 3.1b-3.1e, which contains four atoms and belongs to space group Pmn2$_1$ with a mirror symmetry only ($M_y : y \rightarrow -y$). The lattice parameters are optimized by first-principles density functional theory (DFT) [86, 87] calculations, in good agreement with other theoretical results [70, 88–90]. Figure 3.1b/3.1c and 3.1d/3.1e show the corresponding ferroelectric phase with spontaneous polarization along $-x/\pm x$ and $y/\pm y$, respectively, while a direct comparison of Figure 3.1b/3.1c and Figure 3.1d/3.1e reveals the ferroelastic order with spontaneous strain along $x$ and $y$, respectively. Below we will separately discuss the ferroelastic and ferroelectric order in monolayer MX.

Ferroelastic order in MX monolayers originates from the fact that their centrosymmetric parent phase can undergo spontaneous relaxation along both $x$ and $y$ direction, resulting in two different orientations of spontaneous lattice strain that are perpendicular to each other. For 2D materials, it can be mathematically described by $2 \times 2$ in-plane transformation strain matrix using the paraelastic structure as a reference state. For one of the two ferroelastic ground states of monolayer GeSe with spontaneous tensile strain along $x$, the lattice parameters $a$ and $b$ are 4.26 and 3.98 Å, respectively, which are 4.10 and 4.10 Å in the reference paraelastic phase. According to its space group Pmn2$_1$, two in-plane lattice vectors are perpendicular to each other. Hence, the corresponding $2 \times 2$ in-plane unit cells $H_x$ and $H_{\text{ref}}$ can be expressed by: $H_x = \begin{bmatrix} 4.26 & 0 \\ 0 & 3.98 \end{bmatrix}$ and $H_{\text{ref}} = \begin{bmatrix} 4.10 & 0 \\ 0 & 4.10 \end{bmatrix}$. Transformation strain matrix $\eta_x$ can be defined using the Green-Lagrange strain tensor, $\eta_x = \frac{1}{2} ( [H_{\text{ref}}^{-1}]^T H_a^T H_a H_{\text{ref}}^{-1} - I )$, [76] where $T$ denotes matrix transpose, and $I$ is a $2 \times 2$ identity matrix. For GeSe, $\eta_x = \begin{bmatrix} 0.041 & 0 \\ 0 & -0.027 \end{bmatrix}$, corresponding to 4.1% tensile strain.
along $x$ and 2.7% compressive strain along $y$. Similarly, we obtained transformation strain matrix $\eta_y$ for the other ferroelastic ground state: $\eta_y = \begin{bmatrix} -0.027 & 0 \\ 0 & 0.041 \end{bmatrix}$, indicating a 2.7% compressive strain along $x$ and 4.1% tensile strain along $y$.

Figure 3.2: Ferroelastic order and spontaneous strain in monolayer GeSe. (a) Total energy of monolayer GeSe with different lattice parameters $a$ and $b$ with respect to their ferroelectric ground state. Solid green line: phase boundary between paraelectric and ferroelectric phases. Red dashed line: phase boundary between two ferroelastoelectric phases with spontaneous strain along $x$ and $y$, respectively, where ferroelectric polarization will be along $+x/-x$ and $+y/-y$, respectively. Two purple dots stand for the two ferroelastoelectric ground states at stress-free condition, while the yellow dots stand for the corresponding lattice parameters upon uniaxial tensile strain along $x$ and $y$ axis. The contour colors illustrate the total energy of unit cell relative to the energy of ferroelectric ground states (purple dots). (b) Minimum energy pathway of coherent ferroelastic transition. The gray dots in (b) correspond to the white stars in (a).

To have a better physical picture of the ferroelastic order, we calculate the total energy of fully relaxed monolayer GeSe as function of lattice parameters $a$ and $b$ with respect to the ground-state state energy. The results are shown in Figure 3.2a. Two degenerate ground states are marked by two purple dots characterized by spontaneous strain along $x$ and $y$, respectively. Under relatively large in-plane lattice parameters, two ferroelastic phases are separated by a phase boundary, whereas at smaller lattice parameters they undergo a ferroelectric-to-paraelectric transition. Furthermore, a
generalized solid-state nudged elastic band (NEB) calculation [91,92] was performed to investigate the coherent structural transition from one ferroelastic state to the other. The calculated minimum energy pathway (MEP) is shown in Figure 3.2b and marked by white stars in Figure 3.2a, and the saddle point is labeled by red dot. The small energy barrier of 19 meV/unit cell in monolayer GeSe highlights the possibility of fast switching upon external mechanical stress despite the fact that such switching is generally mediated by domain wall motion. It is worthy to note that the ferroelastic transition does not go through the centrosymmetric paraelectric phase where Ge and Se atoms overlap exactly in the 2D plane. Instead, as shown in Figure 23.2b, the Ge and Se atoms rotate around each other.

3.1.3 Ferroelectric Order and Spontaneous Polarization in Monolayer MX

Ferroelastic order does not necessarily correlate with ferroelectric order. Monolayer black phosphorus is a good example which shares a similar structure with monolayer MX, hence it is a 2D ferroelastic material. Nonetheless, due to single phosphorus element no ferroelectric order is expected. In contrast, the two distinct chemical elements in monolayer MX give rise to appreciable difference in electronegativity and large relative displacement, consequently they are very likely to have large spontaneous polarization. To confirm the above speculation, we need to choose an adiabatic pathway for the ferroelectric transition, and calculate its spontaneous polarization using the Berry phase approach based on the Kohn-Sham wavefunctions from first-principles DFT calculations. First, we scanned the full potential energy surface by fixing the Ge atoms in the unit cell, shifting Se atoms with respect to their hypothetic centrosymmetric position in the $x-y$ plane, and then relaxing the atoms along the $z$-axis. Here, the lattice parameters are constrained to the initial ground state, and the centrosymmetric phase serves as the reference point of total energy. Figure 3.3a clearly shows that two ferroelectric ground states are located at the minima of two potential energy wells. Our NEB calculations show that these two energy minima are connected by two MEPs (marked with green lines) which deviate away from the central paraelectric phase, implying that bulk ferroelectric transition is realized by the relative rotation of Ge and Se atoms around their paraelectric positions, rather than by a straight translation between two ferroelectric
phases (marked as red line). The calculated minimum energy barrier for monolayer MXs spans a wide range, i.e. about 7, 33, 95, and 464 meV/unit cell for SnSe, SnS, GeSe, and GeS, respectively, suggesting a possibility to fine-tune the kinetic barriers of ferroelectric transition via stoichiometric controls of M and X chemical elements for device applications.

Figure 3.3: Ferroelectric order, spontaneous polarization, and elastic strain effect in monolayer GeSe. (a) Potential energy surface with fractional shift of Se atoms along $x$ and $y$ with the fixed ground-state lattice parameters. Green lines: MEP for ferroelectric transition. Red line: adiabatic path for polarization calculation. Two black dots: degenerate ferroelectric ground states. (b) Double-well potential of monolayer GeSe along the adiabatic path. (c) Calculated total polarization as a function of normalized displacement where the centrosymmetric paraelectric phase (0% displacement) is at the center, and two ferroelectric ground states are at two ends. $P_q$ represents the polarization quanta. (d) Effect of uniaxial strain on the total energy and total spontaneous polarization. (e) Spontaneous polarization of ferroelectric GeSe monolayer as function of uniaxial strain. (f) Minimum energy barrier of coherent ferroelectric transition as function of uniaxial strain without the constraint through the paraelectric phase.

The adiabatic path through MXs centrosymmetric paraelectric phase allows us to calculate the total polarization using modern theory of polarization based on the Berry phase approach [33, 39].
Mathematical expression of spontaneous polarization $P_s$ in Wannier representation is given by

$$P_s = P^f - P^i = \frac{1}{V} \sum_j \left( q^f r^j - q^i r^j \right) - \frac{2ie}{(2\pi)^3} \sum_{\text{occ}} \int_{\text{BZ}} \psi_n^f(\mathbf{k}) \mathbf{e}^{-i\mathbf{k} \cdot \mathbf{R}} \left[ \left.u_n^f \right| \frac{\partial u_n^f}{\partial \mathbf{k}} \right] - \left[ \left.u_n^i \right| \frac{\partial u_n^i}{\partial \mathbf{k}} \right] \right], \quad (3.1)$$

where “i” and “f” refer to the initial and final positions, and $u_{nk}(\mathbf{r})$ is the periodic part of Bloch wave functions $u_{nk}(\mathbf{r}) = e^{-i\mathbf{k} \cdot \mathbf{r}} \psi_{nk}(\mathbf{r})$. $V$ is the volume of the unit cell, and the integral is performed over the first Brillouin zone. Here, the initial state is the centrosymmetric structure, and the final state is the polarized one. The above formula has taken into account both ionic and electronic contributions to the total polarization. As shown in Figure 3.3b and Figure 3.3c, monolayer GeSe has a large spontaneous polarization of 357.0 pC/m, rendering an effective bulk polarization of 35.7 $\mu$C/cm$^2$ if we assume an approximate layer thickness of 1 nm. The latter includes both van der Waals distance and the intrinsic thickness of monolayer GeSe. Here, the spontaneous polarization was rigorously determined by identifying the continuous evolution of total polarization as a function of normalized displacement through the centrosymmetric paraelectric state, as shown in Figure 3.3c for monolayer GeSe. The lines with different colors in Figure 3.3c are shifted by multiple polarization quanta $P_q$. Spontaneous polarizations for GeS, SnS, and SnSe monolayers are 484, 260, and 181 pC/m, respectively, corresponding to effective polarization of 48.4, 26.0, and 18.1 $\mu$C/cm$^2$, which agrees well with a very recent work by Fei et al. [93] Monolayer GeS owns the largest spontaneous polarization owing to a sizable electronegativity difference of the two elements and, more importantly, the largest displacement between Ge and S atoms with respect to the paraelectric state.

Furthermore, as shown in Figure 3.3b, the total energy as function of spontaneous polarization along the adabatic path (the red line in Figure 3.3a) exhibits a characteristic double-well potential of ferroelectrics. The solid curve in Figure 3.3b is obtained by fitting the energy as function of polarization in a sixth-order polynomial based on the Landau-Devonshire theory of ferroelectrics [94],

$$E(P) = \frac{1}{2}aP^2 + \frac{1}{4}bP^4 + \frac{1}{6}cP^6,$$

where $E(P)$ is the total energy with respect to the paraelectric phase and $P$ is the total polarization. By including the energy contribution from the conjugate...
external electric field $\mathcal{E}$ and internal polarization, one can define electric enthalpy $F = E - \mathcal{E}P$, from which we can estimate the ideal coercive field $\mathcal{E}_c$ from the maximum slope of $E(P)$, i.e. $\mathcal{E}_c = \max(dE/dP)$, between the minimum and saddle point [84]. The calculated ideal coercive field $\mathcal{E}_c$ for monolayer GeSe is 0.623 V/nm under stress-free condition, which increases as the increasing of uniaxial strain along the polarization direction. However, two issues need to be kept in mind. First, the fitted $E - P$ curve serves as a guidance only, as in principle the fitting parameters $a$, $b$, and $c$ are temperature dependent which is not reflected in the present DFT calculations carried out at 0 K. Second, ideal coercive field $\mathcal{E}_c$ refers to the electric field required to reverse the polarization in a coherent transformation throughout the whole crystal, while in reality true coercive field corresponds to critical electric field that destabilizes the domain walls, which is often much smaller than the ideal coercive field $\mathcal{E}_c$.

2D materials can often sustain large elastic strain [95, 96], making it distinctly different from their 3D bulk counterpart. Specifically, the dual ferroic (ferroelastic and ferroelectric) properties in MX monolayers may be subject to continuous alternations upon elastic stress, which may offer a facile control of charge polarization and phase transition barrier and thus provide a broader design space for optoelectronic devices. To verify this concept, we apply a uniaxial tensile strain along x-axis (i.e. ferroelectric direction) and fully relax the lattice parameter $b$ and the atomic positions. The resulted strain-dependent energy profile as function of total polarization is presented in Figure 3.3d for monolayer GeSe, manifesting a large effect of elastic strain on the spontaneous polarization. For example, by increasing uniaxial strain from 0% to 6%, the spontaneous polarization in monolayer GeSe is markedly enhanced from 357 pC/m to 430 pC/m. The strain-dependent polarization as shown in Figure 3.3e also allows us to estimate the piezoelectric coefficient $e_{11} \equiv (\partial P_x)/(\partial \epsilon_{xx})$ of 1.31 nC/m for monolayer GeSe, which is in good agreement with the recent work by Fei et al. [70]. Furthermore, as shown in Figure 3.3f, the strain-dependent coherent ferroelectric transition energy barrier enlarges from 0.08 to 0.30 eV/unit cell as strain increases from 0% to 6%. The above results unequivocally demonstrate that elastic strain can have significant impact on both spontaneous polarization and coherent transition barriers, highlighting strain
modulation as a potential avenue for fine-tuning ferroelectric properties of materials. To verify the stability of multiferroicity under room temperature, we have performed \textit{ab initio} molecular dynamics (AIMD) simulations for monolayer GeSe at 300K using NPT ensemble at zero pressure and a time step of 0.5 fs. Three configurations at \( t = 0, 900 \) fs, and \( 1800 \) fs. These snapshots of AIMD configurations demonstrate that the multiferroicity is preserved along the initial vertical direction after \( 1800 \) fs. However, it should be noted that the limited AIMD simulations cannot tell the absolute stability and the exact transition temperature of multiferroicity. Moreover, we cannot simply use the energetic difference between ground-state ferroelectric phase and paraelectric phase to determine the transition temperature. Instead, one has to apply the fourth order Landau theory to estimate transition temperature for ferroelectric crystals. This was also pointed out by Fei et al. recently and they applied Monte Carlo simulations with an effective Hamiltonian method and demonstrated robust ferroelectricity in four MX materials beyond room temperature. [93] However, the Curie temperature of the multiferroic transition is more complicated than ferroelectric transition as Curie temperature depends on the mechanical boundary condition and one has to include the coupling of the two order parameters (ferroelectric and ferroelastic). This problem can be addressed using the phase-field approach [97] which can provide a quantitative prediction of both the transition temperature and the domain structure at a given temperature under a stress-free condition. However, this approach will involve substantial development that is beyond the scope of the current work, and is currently under investigation.

3.1.4 Ferroelectric Domain Wall in Monolayer MX

The potential energy surfaces shown above illustrate only the thermodynamic properties of perfect crystalline monolayer MX, whereas in reality ferroelectric transition is governed by domain wall motion, analogous to dislocation motion in solids. It is therefore highly important to explore the energetics and transition pathway of ferroelectric domain wall that are relevant to experimental observation and device characterization. In order to acquire a domain wall configuration, we first construct a supercell of pristine monolayer MX consisting of 24 unit cells repeated along the \( b \) \( (y) \) axis with ferroelectric polarization along \( +x \) direction. We then flip the polarization direction of
the first 12 unit cells, forming a supercell that contains two 180° ferroelectric domain walls: one at the center \((y=b/2)\) and the other at the boundary \((y=0)\). Such configuration obviously satisfies the periodic boundary condition required by first-principles DFT using plane wave basis.

An example of 180° ferroelectric domain wall in monolayer GeS is shown in Figure 3.4a with its polarization direction indicated by blue and purple arrows. The domain wall energies were then calculated by the energy difference between the fully relaxed supercell and pristine monolayer
MX, which yield 116, 56, 24, and 8 meV/Å for GeS, GeSe, SnS, and SnSe, respectively. They correspond to effective domain wall energies of 186, 90, 38, and 13 mJ/m² if we assume a vdW thickness of 1 nm. The effective domain wall energies in monolayer GeS and GeSe are similar to the 180 Å ferroelectric domain wall energies in prototypic PbTiO₃ of 132 and 169 mJ/m² for Pb-centered and Ti-centered domain wall, respectively, while SnS and SnSe have the domain wall energies that are similar to BaTiO₃ with 7.5 and 16.8 mJ/m² for Ba-centered and Ti-centered domain wall, respectively [98]. We then investigate the energy barrier of 180° ferroelectric domain wall migrating along the +y direction using first-principles NEB method. Here we focus on monolayer GeS. Its final configuration is obtained by translating the supercell of the initial configuration by one lattice vector along the +y axis. Total nine images including the initial and final configurations as well as the linearly interpolated intermediate ones were used in the subsequent NEB calculation. The corresponding MEP is shown in Figure 3.4b with the images labeled by 1 to 9. The MEP plot reveals two similar barriers of ~1.6 meV/Å, owing to the fact that the two Ge-S pairs in a single unit cell are related by translational symmetry along the diagonal direction. It is important to notice that the migration barrier of ~1.6 meV/Å is remarkably small, which implies that the ferroelectric domain wall in monolayer GeS is highly mobile. Once nucleated, domain wall assisted switching can proceed very fast. We have also calculated MEPs for the other three MX monolayers which are much smaller than that of GeS, indicating that their domain wall motion can be even much faster. It is worth to mention that we also investigated the 90° domain wall, however, the supercell eventually relaxed back to single ferroelectric phase due to the large elastic strain energy residing in the supercell. Hence, it is more likely to observe the 180° domain wall discussed above.

In paraelectric phase, Ge and S of each local pair exactly overlap on the x-y plane, while in the ferroelectric phase they are shifted with respect to each other. Therefore, one can extract the domain wall width by calculating the relative displacement of each Ge-S local pair in the supercell. Such relative displacement essentially serves as an order parameter which has two components: one along the x direction, and the other along the y direction. Figure 3.4c shows the corresponding
$x$ component of relative displacement as a function of the centers of each Ge-S pair, while the $y$ component is presented in Figure 3.5. Thus, according to Figure 3.4c, the domain wall width of monolayer GeS is about 1 nm which remains almost unchanged for all nine images in the MEP. By comparing the relative M-X displacement for all four materials, we find that GeS has the smallest domain wall width and the largest relative M-X displacement, followed by GeSe, SnS, and SnSe with gradually increased width and reduced displacement, which is consistent with their ground state structure where GeS has the highest spontaneous strain.

Figure 3.5: Ferroelectric domain wall (180°) in monolayer GeS. (a) Minimum energy pathway of domain wall motion with initial, intermediate, and final configurations labeled from 1 to 9. (b) Atomic structures of the initial (#1), middle (#5), and final (#9) configuration. Blue and purple arrows illustrate the polarization direction of the left-hand and right-hand side of domain wall in monolayer GeS, respectively. (c,d) Relative displacement of each Ge-S pair along +x and +y direction, respectively. Colors from blue to red refer to the configurations #1-#9, corresponding to the colors of the dots marked in (a). (e) 2D plots of relative displacement of each Ge-S pair for the configurations #1, #5, and #9 in the supercell. Small blue and purple arrows illustrate the relative displacement direction of each Ge-S pair. Gray spheres stand for the centers of each Ge-S pair.
The 180° ferroelectric domain wall in monolayer MX is essentially a one-dimensional interface between two ferroelectric phases with antiparallel electric polarizations. It is inevitably accompanied by distinct electronic structure localized in the vicinity of the domain wall, which is worth of a detailed study. Figure 3.4d presents the DFT-PBE band structure of monolayer GeS supercell with 180° ferroelectric domain wall, where different colors indicate the relative localization of electronic states around the domain wall: purple for the states near the domain wall, and cyan for the states away from the domain wall. It clearly shows that the four lowest conduction bands (from the conduction band minimum (CBM) to CBM+3) reside near the domain wall, forming two sets of degenerate bands, i.e. degenerate CBM and CBM+1, and degenerate CBM+2 and CBM+3. Figure 3.4e shows the corresponding real-space wavefunctions for CBM at Γ point, while the wavefunction at CBM+2 can be found in Figure 3.6. The localized electronic wavefunctions positioned at the low energy conduction bands suggest that they may be easily detected via electronic and optical measurement, for examples, by scanning tunneling microscopy and photoluminescence measurement.

Figure 3.6: Electronic structures of monolayer GeS with 180° domain wall. (a) Electronic band structure. (b) Top and front view of the Kohn-Sham wavefunction at the conduction band minimum (CBM). (c) Top and front view of the Kohn-Sham wavefunction at the CBM+2. Domain wall is illustrated by dashed blue line. Blue and purple arrows indicate the polarization direction on the left and right hand side of the domain wall.
3.1.5 Electronic Structure of Ferroelectric Monolayer MX

Figure 3.7: Electronic band structures of monolayer GeS, GeSe, SnS, and SnSe. Blue lines: calculations using DFT-PBE with spin-orbit coupling included. Red lines: calculations using many-body perturbation theory within the $G_0W_0$ approximation with the spin-orbit coupling included as well.
The ferroelectricity and ferroelasticity of monolayer MX discussed above will have more profound impact if they are also coupled to their electronic and optical properties. Figure 3.8a and Figure 3.7 show the calculated quasiparticle band structure of MX monolayers where both spin-orbit coupling and quasiparticle effect are included. The quasiparticle effect was taken care by many-body perturbation theory within the $G_0W_0$ approximation [99,100]. The results demonstrate that MX monolayers are intrinsically 2D semiconducting materials with indirect band gaps ranging from 2.6 eV to 1.1 eV. We also calculated their photoabsorption spectra by solving two-particle Bethe-Salpeter equation [101,102] based on the quasiparticle energies and the screened Coulomb interactions obtained from the GW calculations. The calculated lowest exciton energies are 2.3, 1.2, 1.8, 1.0 eV for GeS, GeSe, SnS, and SnSe, respectively, whereas their corresponding direct quasiparticle transition energies are 2.9, 1.6, 2.3, and 1.3 eV, resulting in large exciton binding energies of 0.6, 0.4, 0.5, and 0.3 eV that are similar to monolayer MoS$_2$ etc [96]. In general, the dimensionality reduction leads to a reduced screening in 2D semiconductors, and consequently the effective Coulomb interaction of electron-hole pair becomes much stronger, hence large exciton binding energies in 2D semiconductors are expected upon photoexcitation. Figure 3.8b presents the calculated imaginary dielectric functions with both excitonic and spin-orbit coupling effect included. It illustrates a highly anisotropic photoabsorption in monolayer GeSe with two giant low-energy excitonic peaks: one at 1.2 eV for the $xx$ component and the other at 1.9 eV for the $yy$ component of the dielectric tensor. Figure 3.8c shows the calculated optical absorbance in percentage [103]. It reveals a strong excitonic optical absorption by monolayer GeSe of 9.4% at 1.3 eV and 5.5% at 2.0 eV corresponding to exciton #1 and #2, respectively. Our results on excitonic behaviors and strong optical absorption agree well with two recent theoretical works [88,89]. Such anisotropic electronic and optical properties are the direct consequences of broken crystalline symmetry due to the ferroelastic spontaneous strain and ferroelectric spontaneous polarization in MX monolayers.
Figure 3.8: Electronic structure of monolayer GeSe. (a) Band structure of pristine monolayer GeSe. (b) Imaginary dielectric function along the $x$ (red) and $y$ (blue) direction. Two arrows in the band structure (a) indicate the quasiparticle gap that corresponds to two excitonic peaks in the imaginary dielectric function (b). (c) Absorbance (%) of monolayer GeSe.

3.1.6 Device Concepts Based On 2D Multiferroic Monolayer MX

The coupled ferroelastic and ferroelectric orders and the polarization-dependent optical properties within the visible range imply their great potentials for 2D mechano-opto-electronic device applications by manipulating electrical, mechanical, and optical external fields to control the internal responses. For example, as shown in Figure 3.9a, by applying large external electric field or bias voltage one can switch the polarization direction along the same axis, and by measuring the current-voltage curve under small bias one can detect the state without any detrimental effect, thus implementing the “write” and “read” functions of 2D ferroelectric memory. The corresponding state can also be detected or “read” by measuring the photocurrent upon photo-illumination since the internal electric field will determine the drifting directions of photoexcited charge carriers. Furthermore, owing to their ferroelastic nature, applied mechanical stress can also control the polarization state along different in-plane axes (Figure 3.9b), which can be “read” by electrical measurement or by polarization-dependent photoluminescence measurement, rendering a 2D ferroelastic memory. Going one step further, one may apply intense coherent light to switch (write) between the two ferroelastic states through optically-controlled domain wall motion [104],
and “read” out the states through the polarization-dependent photoluminescence measurement, enabling a 2D ferroelastoelectric non-volatile photonic memory (Figure 3.9c). Remarkably, this optical induced domain wall motion has been observed very recently by Rubio-Marcos et al. in BaTiO$_3$ single crystal. [105] Although the present material is different from BaTiO$_3$, it is plausible to observe the similar phenomena in MX monolayers due to the following reasons. First, the domain wall in BaTiO$_3$ possesses two-dimensional nature, while it is 1D in MX monolayers. This means, the optical power needed to initiate the domain wall and flip the ferroelectric polarization will be proportional to domain wall length in MX monolayers, while it is proportional to domain wall area in BaTiO$_3$, thus the scaling for ferroelectric transition is more preferable in 2D MX monolayers. Second, the calculated domain wall energy and migration barrier show large variations among different MX monolayers, thus it offers a wide range of materials choices to achieve optical induced transition with less power. It is important to point out that SnTe, a cousin of the present four MX monolayers, has been synthesized and demonstrated with 2D ferroelectricity very recently by Chang et al. [106] Thus the present device proposal can be tested in 2D SnTe. Finally, as shown in Figure 3.9d, MX monolayers may serve as an ideal materials platform for realizing 2D ferroelectric photovoltaics, because (i) their optical bandgap falls within the visible range from 1.0 eV to 2.3 eV; (ii) their optical absorption is strong due to the pronounced excitonic effect common among different MX monolayers, thus it offers a wide range of materials choices to achieve optical induced transition with less power. It is important to point out that SnTe, a cousin of the present four MX monolayers, has been synthesized and demonstrated with 2D ferroelectricity very recently by Chang et al. [106] Thus the present device proposal can be tested in 2D SnTe. Finally, as shown in Figure 3.9d, MX monolayers may serve as an ideal materials platform for realizing 2D ferroelectric photovoltaics, because (i) their optical bandgap falls within the visible range from 1.0 eV to 2.3 eV; (ii) their optical absorption is strong due to the pronounced excitonic effect commonly shared by two-dimensional semiconducting materials (Figure 3.8b and Figure 3.8c, where monolayer GeSe can absorb more than 6% upon photo-illumination); and (iii) the internal in-plane electric field in ferroelectric MX can greatly enhance the separation of the photo-excited charge carriers to form photocurrent. However, we would like to emphasize that all these exciting proposals are at their conceptual stage, and experimental demonstrations have to be carried out in future to validate these concepts in these monolayer MXs. It is worth to emphasize that the ultrathin two-dimensional nature has another significant implication, that is, the required energy consumption for the “write” and “read” operations in the above 2D ferroelectric/ferroelastic/photonic memory will be much smaller than that in thick films. Since the kinetic barrier of the domain wall motion which controls the switching process is small, the operational speed should be very fast. In addition,
the non-volatile photonic memory proposed here is based on the ferroelastic transition between two symmetrically-equivalent crystalline states, which is distinct from the recently-demonstrated first photonic switch based on photo-induced crystalline-to-amorphous phase transition. [107] It is worth to note that, upon the completion of this work, we learned of two independent works including an experimental study of ferroelectricity in 2D SnTe by Chang et al. [106] and a theoretical study of four MX materials by Wu et al. [90] but with different emphasis between the present work and the above two.

To summarize, in this work we discussed both ferroelasticity and ferroelectricity as well as the electronic, optical, and domain wall properties of monolayer Group IV monochalcogenides. Monolayer MX represents a class of 2D multiferroic semiconductor with large in-plane spontaneous polarization, spontaneous lattice strain, and small domain wall energy, an important addi-
tion to the existing realm of multiferroic bulk materials, interface structures, and thin films. The calculated energy barriers of coherent ferroelectric-to-paraelectric and ferroelastic-to-paraelastic transitions show that their dual ferroic orders are thermodynamically stable at room temperature. In addition, the ferroelectric polarization and coherent transition barrier can be easily tuned via elastic strain engineering. The in-plane ferroelectric polarization in 2D MX monolayers makes them distinctly different from ferroelectric thin-films such as perovskite compounds with out-of-plane polarization whose thickness is intrinsically limited by the amplified out-of-plane depolarization effect [85]. Conversely, a minimal in-plane feature size is expected in MX monolayers due to in-plane depolarization. The strongly-coupled ferroelastic and ferroelectric orders and the polarization-dependent excitonic absorption and photoluminescence within the visible spectrum may enable potential applications for 2D mechano-opto-electronic applications such as 2D ferroelectric excitonic photovoltaics and 2D memory devices including ferroelectric, ferroelastic, and ferroelastoelectric non-volatile photonic memory. The present findings will open up new avenues for miniaturized low-power optoelectronic and photonic applications of 2D multiferroic materials with coupled electronic, optical, mechanical, and even magnetic properties if the symmetry-breaking induced valley-dependent polarization is also considered.
3.2 Van der Waals Stacking Induced Topological Phase Transition in Layered Ternary Transition Metal Chalcogenides

Novel materials with nontrivial electronic and photonic band topology are crucial for realizing novel devices with low power consumption and heat dissipation, and quantum computing free of decoherence. Here we theoretically predict a novel class of ternary transition metal chalcogenides that exhibit dual topological characteristics: quantum spin Hall insulators (QSHIs) in their 2D monolayers and topological Weyl semimetals in their 3D noncentrosymmetric crystals upon van der Waals (vdW) stacking. Remarkably, we find that one can create and annihilate Weyl fermions, and realize the transition between Type-I and Type-II Weyl fermions by tuning vdW interlayer spacing, based on which we provide the missing physical picture of the evolution from 2D QSHIs to 3D Weyl semimetals. Our results also show that these materials possess excellent thermodynamic stability and weak interlayer binding and some of them were synthesized two decades ago, implying their great potentials for experimental synthesis, characterization, and vdW heterostacking. Moreover, their ternary nature will offer more tunability for electronic structure by controlling different stoichiometry and valence charges. Our findings provide an ideal materials platform for realizing QSH effect and exploring fundamental topological phase transition, and will open up a variety of new opportunities for two-dimensional materials and topological materials research.

3.2.1 Brief Introduction to Topological Materials

The seminal discovery of quantum spin Hall (QSH) effect [108–113] engendered a new chapter of topological materials research in condensed matter physics and materials science [64, 65], followed by the discoveries of three-dimensional (3D) topological insulator [114–118], quantum anomalous Hall insulator [119, 120], topological crystalline materials [121–126], Weyl semimetals [127–136] etc. A generic aspect is the presence of special surface/edge states that are topologically protected upon small perturbations, hence inelastic scattering induced heat dissipation is

†Reprinted with permission from “Van der Waals Stacking Induced Topological Phase Transition in Layered Ternary Transition Metal Chalcogenides” by Junwei Liu, Hua Wang, Chen Fang, Liang Fu, and Xiaofeng Qian, Nano Letters 17, 467-475 (2017). Copyright © 2017, American Chemical Society.
minimized. In contrast, conventional electronics suffers from severe local heating as any structural or chemical defect could introduce additional scattering and reduce carrier transmission. These topological phases, if materialized and integrated at the device level, could be advantageous for many novel low-power and low-dissipation electronic applications. Along with materials discoveries with individual topological phases, there is an increasing interest to understand topological phase transition, for examples, from nontrivial to trivial topology and from one nontrivial phase to another [64, 65] through chemical doping, elastic strain engineering, electric and magnetic field perturbations, etc. In parallel, 2D transition metal dichalcogenides (TMDCs) also arouse substantial interest [60, 62, 77] in their distinct properties, for example, extraordinarily enhanced excitonic photoluminescence in noncentrosymmetric monolayer 1H-MoS\(_2\) that consists of three planes of 2D hexagonally packed atoms, S-Mo-S, in a Bernal ABA stacking. [137] Most TMDCs have trivial topology, either metallic or semiconducting with large gap. However, different from their polymorphic 1H variant, 1T structure of binary TMDCs such as MoTe\(_2\) and WTe\(_2\) possess centrosymmetric characteristic with three atomic planes stacked in a rhombohedral ABC sequence. 1T-MoTe\(_2\) and WTe\(_2\) are subject to the Peierls instability, resulting in a dimerized 1T’ structure in their monolayers, which surprisingly leads to a variety of novel quantum phenomena such as nontrivial electronic topology [74, 138] and nonsaturating giant magnetoresistance [139]. In particular, their monolayers were predicted to exhibit QSH effect,36 and the weak vdW coupling between different monolayers makes them suitable to realize multi-functional materials/devices by stacking with other 2D materials, [61] such as topological field transistors. [74] Encouragingly, there are great experimental progresses of nontrivial 2D TMDCs. For example, Peierls-distorted MoTe\(_2\), though metastable, has been successfully synthesized in experiments, and recently characterized with a fundamental gap of about 60 meV by both optical and electrical measurements. [140–142] Recently Fei et al. provided some experimental evidence for the QSH effect in monolayer WTe\(_2\). [143] Nevertheless, special encapsulation and assembly are required to prevent them from transforming into the more stable but topologically trivial 1H phase. Meanwhile, there are limited choices within binary TMDCs for the discovery and design of materials with nontrivial topology. Therefore, for the
ease of experimental realization and device implementation, it is highly desirable to go beyond the present binary MX$_2$ materials and explore other possibilities for large gap topological materials that are thermodynamically stable.

### 3.2.2 Prediction of New Class of Topological Materials

Here we predict four layered ternary transition metal chalcogenides (TTMC), namely MM'Te$_4$, with M=(Nb, Ta) and M'=(Ir, Rh), [144,145] as a new class of topological materials which displays QSH effect with sizable fundamental gaps in their monolayers, become topological Weyl semimetals in their bulk form. Remarkably, the vdW interlayer interaction drives topological phase transi-

---

Figure 3.10: Atomic configuration of ternary transition metal chalcogenides MM'Te$_4$. (a) Top and side view of two-dimensional MM'Te$_4$ in monolayer 1T'. Dashed purple lines indicate in-plane zig-zag metallic bond along the x direction owing to Peierls distortion in the y direction. (b) Side view of three-dimensional MM'Te$_4$ in their bulk T$_d$ structure. Both M and M' are transition metal atoms. M, M', and Te are indicated by blue, orange, and gray atoms. M stands for Ta or Nb, M' stands for Ir or Rh in the present work.
tion via formation and splitting of Dirac points and the latter results in the emergence of topological Weyl nodes and the smooth transition between Type-I and Type-II Weyl fermions. [138, 146] All four ternary compounds are found to be thermodynamically and dynamically stable, and their interlayer binding energies are slightly weaker than that of graphite, making them promising for experimental synthesis, mechanical exfoliation, vdW heterostructuring, and direct electronic characterization thereafter, thereby providing an ideal materials platform for exploring both novel device applications and fundamentals of topological phase transitions.

Bulk ternary transition metal chalcogenides MM’Te₄ are noncentrosymmetric orthorhombic layered materials [144, 145] with space group Pmn2₁ (commonly known as T₉ polymorph) as shown in Figure 3.10, where M=(Nb, Ta) of Group 5, and M’=(Ir, Rh) of Group 9. Different from their bulk, monolayers MM’Te₄ are centrosymmetric 2D semiconductors with space group P2₁/m (known as 1T’ polymorph) whose inversion center is located at the middle of two adjacent M atoms or M’ atoms along the dimerization direction (+y). The atomic structures were fully relaxed using first-principles density functional theory (DFT) with interlayer vdW interactions taken into account. Taking NbRhTe₄ as an example, each layer of bulk T₉-NbRhTe₄ is similar to monolayer 1T’ TMDCs, i.e. exhibiting Peierls distortion induced dimerization along y with zig-zag quasi-one dimensional chains formed along x. Compared to the 1T’ TMDCs, the unit cell along the dimerization direction doubles as the two metallic elements repeat as [–M’-M–M-M’–]. Here, single and double dashes denote short and long bond between adjacent metallic atoms. This doubling is due to different electronegativities of M and M’, resulting in charge transfer between M and M’, hence stronger bonding and shorter bond length in M-M’ than M–M or M’–M’ bond. Consequently, it is energetically more stable than the structure with direct [M’M] repeating units.
Figure 3.11: Electronic band structures of 1T’-MM’Te$_4$ with and without spin-orbit coupling by DFT calculations with GGA and hybrid HSE06 exchange-correlation functional.

Ground-state atomic structures of monolayer 1T’ and bulk T$_d$-MM’Te$_4$ were calculated by first-principles density functional theory (DFT) [87] implemented in the Vienna Ab initio Simulation Package (VASP) with plane wave basis [147] and the projector-augmented wave method [148].
We adopted the Perdew-Berke-Ernzerhof (PBE)’s form [149] of exchange-correlation functional within the generalized-gradient approximation (GGA) [150] and the Monkhorst-Pack k-point sampling for the zone integration. More specifically, an energy cutoff of 400 eV and Monkhorst-Pack k-point sampling of 12×3×3 for bulk and 12×3×1 for monolayer were chosen to fully relax the atomic structures with maximum residual force less than ~0.01 eV/Å. The interlayer van der Waals (vdW) interactions were taken into account by applying the optB88-vdW correlation functional proposed by Klime et al. [151] To confirm the dynamical stability of the relaxed structures, phonon band structures were calculated by using density-functional perturbation theory. [152]

To investigate the topological nature of MM’Te$_4$, we first transformed the Kohn-Sham eigenstates of DFT calculations into a set of highly localized quasitatomic orbitals accompanied with first-principles tight-binding Hamiltonian. [153, 154] The latter was used for the calculation of topological indices such as $Z_2$ topological index and Chern number. More specifically, we computed the $Z_2$ topological index of monolayer materials using two independent methods, including the parity product at time-reversal invariant momenta [155] and the Berry-phase based $n$-field method. [156] For bulk T$_d$-MM’Te$_4$, we first performed a coarse scan to locate the rough regions of potential Weyl points between the top valence band and the bottom conduction band, and then used simplex search method [157] to find their exact locations with an energy tolerance of 10$^{-10}$ eV. Subsequently, we constructed a triangulated sphere near each Weyl point and calculated the corresponding Berry curvature and Chern number. Here the radius of sphere was chosen to be small enough to avoid two Weyl points within a single sphere. Bulk, edge and surface DOS are calculated using the imaginary part of the retarded Green’s function [158] of semi-infinite atomistic models of 2D and 3D MM’Te$_4$. A kinetic energy cutoff of 300 eV and a Monkhorst-Pack k-point sampling of 8×3×3 for bulk MM’Te$_4$ and 8×3×1 for monolayer MM’Te$_4$ were used for the above electronic structure calculations. As GGA functionals usually underestimate bandgaps, we also performed DFT calculations with hybrid HSE06 exchange-correlation functional [159] for 1T’-MM’Te4, and except TaIrTe$_4$ which is at the border line, all other three materials maintain their nontrivial $Z_2$ topology. However, hybrid functional calculations for bulk T$_d$ structures
become extremely heavy, therefore in order to be consistent we present the results calculated by DFT with GGA functionals for both monolayer 1T' and bulk Td MM'Te. The detailed electronic structures with hybrid functional are shown in Figure 3.11.

Figure 3.12: Electronic structure of quantum spin Hall insulators 1T'-MM'Te4 using monolayer 1T'-NbRhTe4 as a representative case. (a) 2D Brillouin zone of monolayer 1T' NbRhTe4 and its projections onto the $k_x$ and $k_y$ edges. Dirac points “Λ” in the absence of spin-orbit coupling are indicated by red dots along the YRY zone boundary. (b) Electronic band structure of monolayer NbRhTe4 with and without spin-orbit coupling indicated by red and blue lines. (c, d) Bulk DOS, edge DOS, and spin ($S_z$)-resolved edge DOS of monolayer NbRhTe4 along the $k_y$ and $k_x$ edge, respectively. Red and blue color in the edge spin DOS indicate the positive and negative spin, respectively.

Without SOC two massless Dirac cones (Λ) show up along the RYR high symmetry line on the Brillouin zone edge indicated in Figures 3.12a and 3.12b, which implies the presence of band inversion around Y. Once SOC is turned on, a band gap of 65 meV is opened up at the Dirac points...
while time reversal symmetry (TRS) is still preserved. The corresponding band structures for all four monolayer MM'Te$_4$ are shown in 3.11. Band inversion at $Y$ together with TRS renders the characteristic of quantum spin Hall insulator. To verify its topology, we calculate the $Z_2$ topological index using two independent methods including the parity product at time-reversal invariant momenta [155] and the Berry-phase based $n$-field method [156], and both methods give $Z_2$ index of 1, thereby confirming the nontrivial topology. The nontrivial topology is also originated from the Peierls distortion in 1T$'$ structure which is indeed very similar to 1T$'$ binary TMDCs, while the microscopic mechanism of band inversion is a little different. Here both the top valence band and bottom conduction band near the Fermi level of monolayer NbRhTe$_4$ are dominated by $d_{x^2-y^2}$, $d_{yz}$, and $d_{z^2}$ of Nb atoms, corresponding to a $d$-$d$ band inversion that is different from the $p$-$d$ band inversion in 1T$'$ binary TMDCs. Furthermore, as all Nb atoms are located in the middle of the three atomic layers, their $Z_2$ topology should be more robust against local environment, such as vdW heterostacking [57, 61]. Given a piece of topologically nontrivial and large enough 2D monolayer MM'Te$_4$ whose edge is interfaced with a trivial material such as vacuum, there will be an unavoidable nontrivial-to-trivial topological phase change across the interface. It guarantees the appearance of two counter-propagating helical states along the edges of 2D host materials in the absence of magnetic field or non-TRS breaking impurities, connecting the bulk valence states to the conduction states. To investigate the edge states, we constructed first-principles tight-binding (TB) Hamiltonians for each monolayer MM’TTe$_4$ and applied Green’s function method to calculate the electronic structure of semi-infinite monolayer MM’TTe$_4$ from which both bulk density of states (DOS) and edge DOS are extracted. Figures 3.12c and d present the corresponding bulk and edge DOS as well as spin ($S_z$)-resolved edge DOS for the $y$ and $x$ edges of monolayer NbRhTe$_4$, respectively. Both unambiguously reveal the fully gapped bulk electronic band structures and the presence of nontrivial edge states. Results for the other three monolayer MM’TTe$_4$ are shown in Figure 3.13. The spin-resolved edge DOS clearly demonstrates the opposite spin carried by two counter-propagating states which cross each other at time-invariant momenta of 1D Brillouin zone, that is, $\Gamma$ and $\bar{\Gamma}$ along the $y$ edge and $\bar{\Gamma}$ and $\bar{X}$ along the $x$ edge guaranteed by TRS.
Figure 3.13: Bulk DOS, edge DOS, and total bulk and edge DOS of four monolayer 1T’-MM’Te₄. Left panels: along the Y. Right panels: along X edges.
Bulk T\textsubscript{d}-MM’Te\textsubscript{4} are formed by vdW bonded MM’Te\textsubscript{4} layers. In principle, if the interlayer interaction is weak, directly stacked multilayer MM’Te\textsubscript{4} could exhibit even-odd oscillation in their Z\textsubscript{2} topological index, and essentially form 3D weak topological insulators. However, the interaction between layers is not as weak as it appears, and the stacking geometry is not as simple as the AA stacking which is obvious according to the symmetry operation of their Pmn2\textsubscript{1} space group.

The absence of inversion symmetry together with the non-negligible interlayer interaction introduces a topological phase transition of bulk T\textsubscript{d}-MM’Te\textsubscript{4} into Weyl semimetal phase. Weyl points are essentially “magnetic monopoles” with topological charges defined by Chern number, i.e., integration of Berry curvature enclosing a single Weyl point divided by 2\pi. It can be either +1 or -1, corresponding to the outgoing and incoming effective magnetic flux, respectively. We performed a detailed scan of bulk electronic band structure to locate the intersecting points between the top valence band and the bottom conduction band, and found all four materials possess at least four Weyl points. Here, as illustrated in Figure 3.14a, bulk T\textsubscript{d}-NbRhTe\textsubscript{4} has eight Weyl points, which can be divided into two groups with their energy-momentum positions \((E, k_x, k_y, k_z)\) of \((0.125 \text{ eV}, \pm 0.16, \pm 0.12, 0)\) and \((-0.066 \text{ eV}, \pm 0.01, \pm 0.07, 0)\), respectively, where momenta are given in the fractional reciprocal lattice coordinate and the two points in the first quadrant are denoted by \(W_1\) and \(W_2\). All the Weyl points are locked on the \(k_z=0\) plane, which is due to their intrinsic Pmn2\textsubscript{1} space group as discussed later. Moreover, all of them are well separated in the momentum space, making it much easier for experimental detection. Correspondingly, the band structures near \(W_1\) and \(W_2\) are plotted in Figures 3.14b and 3.14c. Constant energy planes of 0.125 eV and -0.066 eV cut through the top valence and bottom conduction bands of \(W_1\) and \(W_2\), respectively, forming electron and hole pockets, which indicates they are Type-II points [138]. The calculated Chern numbers enclosing \(W_1\) and \(W_2\) are -1 and +1, implying negative and positive “magnetic monopole”, i.e., effective net Berry flux going inward and outward, respectively. The corresponding Berry curvatures are shown in Figures 3.14c and 3.14e, which are highly anisotropic due to the vdW stacking nature of bulk T\textsubscript{d}-MM’Te\textsubscript{4}. 

63
Figure 3.14: Electronic structure of Weyl semimetal bulk $T_d$-NbRhTe$_4$. (a) 3D Brillouin zone of Weyl semimetal bulk $T_d$-NbRhTe$_4$ and its projections onto top and bottom surfaces along the $z$ direction. Eight gapless Weyl points are located in the $k_z=0$ plane, indicated by red and blue dots (i.e. Chern number of +1 and -1). “$W_1$” and “$W_2$” represent the two Weyl points in the first quadrant of $k_z=0$ plane. Their projections on top and bottom (001) surfaces form Fermi arcs passing through individual pairs of Weyl points. (b,c) Band structure close to “$W_1$” and “$W_2$”, respectively. Conduction and valence bands intersect with a constant energy plane (indicated by a gray mesh), forming a linear cone and separating electron and hole pockets. (d,e) Berry curvature $\Omega$ around Weyl points $W_2$ and $W_2$ with high anisotropy, respectively.

A unique feature of Weyl semimetals is the presence of exotic Fermi arcs on surfaces, terminated at the projections of a pair of Weyl points with opposite Chern numbers, which is indeed confirmed in our calculations as shown in Figure 3.15. All energy contours of the bulk states (Figures 3.15a, 3.15b, 3.15e and 3.15f) are normal closed loops, while for the surface states (Figures 3.15c, 3.15d, 3.15g and 3.15h) there are some nontrivial open Fermi arcs ending exactly at the surface projections of the bulk Weyl points with opposite Chern numbers marked by red and blue circles. Moreover, some Fermi arcs spread a very large region of surface BZ since the Weyl points are very well separated, which makes experimental detection such as angle-resolved photoemission spectroscopy much easier.
Figure 3.15: Bulk and surface density of states of bulk T$_d$-NbRhTe$_4$. (a,e) Bulk DOS at the two energy levels of 0.125 eV and -0.066 eV corresponding to Weyl points $W_1$ and $W_2$, respectively. (b,f) Bulk DOS of the local region around the two Weyl points in the first and fourth quadrants (marked by green dashed box in a and e). (c,g) Surface DOS on the top surface of NbRhTe$_4$. (d,h) Surface DOS on the bottom surface of NbRhTe$_4$. Weyl points are marked by red and blue dots on each figure.

The above results demonstrate that monolayer 1T$'$-MMTe$_4$ is a 2D TI, and bulk T$_d$-MM$'$Te$_4$ is a Weyl semimetal. Since the former is the decoupled layers limit of the latter with AA stacking, it is important to understand the evolution of topological phase transition between these two cases as the layer separation decreases. Direct first-principles calculations will be too heavy. Nonetheless, this can be achieved by using a linear interpolation between two effective tight-binding Hamiltonians, one with a large lattice constant $c$ (18.46Å used here for T$_d$-NbRhTe$_4$) corresponding to the decoupled layers limit and one with the bulk T$_d$ structure ($c=13.46$Å). The linear interpolation simulates the evolution of the layer separation. Subsequently, one can determine the whole evolution process of Weyl points by scanning the $k_z = 0$ plane of Brillouin zone, locating the Weyl points.
if any, and verifying their topology by computing the associated Chern numbers. Physically, the interlayer coupling plays two roles: (i) it breaks the inversion symmetry and splits the spin degeneracy of all bands and (ii) it modifies the band gap near $X$ significantly and causes band inversions nearby. Each inversion of two non-degenerate bands in a time-reversal invariant system leads to the creation or annihilation of four Weyl points [127].

Figure 3.16: Topological phase transition from stacked QSHIs to Weyl semimetal. (a) Topological evolution of Weyl points locked on the $k_z=0$ plane with decreasing effective interlayer distance in bulk $T_d$-NbRhTe$_4$. Solid dots: Chern number of +1; empty dots: Chern number of -1. Type I and Type II Weyl points are labeled at the beginning and the end of the evolution trajectories in the first quadrant. Final eight Weyl points at equilibrium $c$ are marked by dark asterisk symbols "*". Four groups of Weyl points are indicated by 1-4, respectively. (b-e) Creation, evolution, and annihilation of the first group of Weyl points in Figure 3.15a as interlayer distance decreases captured by a simple $k \cdot p$ model. Three ellipses correspond to $E_{1+}$ (blue), $E_{1-}$ (red) and $E_2$ (green), as $m_{1,2}$ increases.

Figure 3.16a shows the results of topological phase transitions for $T_d$-MM'Me$_4$ with lattice constant $c$ approaching the equilibrium value, with color changing from purple to blue as $c$ decreases. Four groups of Weyl points are clearly seen in Figure 3.16a. The first group of Weyl points initially
appear at relatively large lattice constant (purple-to-blue dots) near \(Y\) when two Dirac points on \(Y\Gamma Y\) are split into four Weyl points which quickly evolve towards \(YR\), and then annihilate with decreasing \(c\). The second group also appears on \(\Gamma Y\) but close to \(\Gamma\), and eventually annihilate on \(\Gamma X\) as well. The third group emerges on \(YR\) close to the original two massive Dirac points of monolayer 1T'-MMTe_{4}, then moves towards \(\Gamma Y\), but they do not reach any high-symmetry line at the equilibrium \(c\), leaving four Weyl points, where the one in the first quadrant has +1 Chern number. The last group emerges on \(\Gamma X\), then moves away from \(\Gamma X\), leaving four Weyl points at the equilibrium lattice constant \(c\), where the one in the first quadrant has -1 Chern number. As a result, there are total eight Weyl points left in the BZ, exactly the same as the ones illustrated in Figure 3.14a for the relaxed bulk T\(_d\)-MM'Te\(_4\). It is worth noting that during the above topological evolution, both Type I and Type II Weyl points show up which are marked in the first quadrant of Figure 3.16a. Moreover, the third group starts with Type I Weyl points, but ends up with Type II Weyl points at equilibrium \(c\), indicating a dramatic band structure evolution as well. This also implies one can induce the transition between Type I and Type II Weyl points by external perturbations such as elastic strain.

### 3.2.3 \(k\cdot p\) Model

This process involves a series of topological phases transitions sharing some common features summarized here. First, all the Weyl points are pinned to the \(k_z=0\) plane due to the symmetry \(\mathcal{T} \ast \mathcal{M}_z \ast \mathcal{G}_y\). Second, two Weyl points emerge from some high-symmetry line and move towards another, perpendicular high-symmetry line. Third, the layer coupling in equilibrium T\(_d\)-MM'Te\(_4\) is not very small, as some Weyl points traverse almost half of the BZ to meet and annihilate. Finally, each creation or annihilation of a pair of Weyl points along a high-symmetry line is accompanied by the change of the topological \(Z_2\)-index on the plane in the BZ spanned by that line and \(k_z\)-axis. Due to the third point above, one does not expect to recover the whole process using an effective \(k\cdot p\) model based at \(Y\). However, below we show that the evolution (i.e. creation and annihilation) of the first group of Weyl points in Figure 3.16a, i.e., when \(c\) is still large, can be captured by a simple model. We start from the monolayer 1T'-MM'Te\(_4\) where we have mirror symmetry \(\mathcal{M}(x, y, z) \rightarrow x, y, -z\) and \(\mathcal{T} \ast \mathcal{M}_z \ast \mathcal{G}_y\).
\((-x, y, z)\), inversion symmetry \(\mathcal{I}(x, y, z) \rightarrow (-x, -y, -z)\) and time-reversal symmetry. One can build the following kp model under these symmetry constraints near \(Y\) where the minimal gap is:

\[
H_{2D}(q_x, q_y) = (\Delta - \epsilon(q))\tau_z + v_y\tau_y q_y + v_x\tau_x s_y q_x,
\]

(3.2)

and

\[
\mathcal{M} = is_x, \\
\mathcal{I} = \tau_z,
\]

(3.3)

\[
\mathcal{T} = K(is_y),
\]

where \(q = k - Y\) is the momentum relative to \(Y\). Here \(\epsilon(q)\) is restricted to an even function of \(q_x\) and \(q_y\), or \(\epsilon(q^2_x, q^2_y)\). We then consider the coupling between the two layers, including only mass terms (\(q\)-independent terms) for simplicity. In bulk \(T_d\text{-MM}^\prime\text{Te}_4\), the inversion symmetry is broken by the interlayer coupling, but we have a new glide plane, \(G : (x, y, z) \rightarrow (x + 1/2, -y, z + 1/2)\). The glide plane operation interchanges the layer index and acts on the spin as a \(\pi\)-rotation about the \(y\)-axis. So we choose:

\[
G = i\sigma_x s_y.
\]

(3.4)

In this basis, the 3D Hamiltonian without interlayer coupling takes the form

\[
H_{3D}(q_x, q_y) = \left[\Delta - \epsilon(q^2_x, q^2_y)\right]\tau_z + v_y\sigma_z\tau_y q_y + v_x\tau_x s_y q_x,
\]

(3.5)

where \(\sigma_i\) acts on the interlayer index within one unit cell. Note that due to the glide plane symmetry, the velocity along the \(y\)-axis has opposite signs for the two layers. The mass terms that couple the two layers should commute with all symmetry operations. There are several mass terms allowed, but we find the following two mass terms important for realizing the evolution of the Weyl points in our numerics:

\[
H_M = m_1\sigma_y s_x + m_2\sigma_z\tau_y s_x.
\]

(3.6)
It is worth commenting that while $m_1$-term is an interlayer coupling, it does not break the inversion symmetry, while the intralayer term $m_2$ breaks the inversion symmetry. As both terms come from layer stacking, we expect both to increase as lattice constant $c$ decreases. The dispersion of the full Hamiltonian $H = H_{3D} + H_M$ is

$$E(q_x, q_y, q_z = 0) = \sqrt{(\epsilon(q_x^2, q_y^2) - \Delta \pm m_1)^2 + \left(\sqrt{v_x^2 q_x^2 + v_y^2 q_y^2 \pm m_2}\right)^2}.$$ (3.7)

The band crossing happens when either

$$\epsilon(q_x^2, q_y^2) = \Delta \pm m_1,$$ (3.8)
$$v_x^2 q_x^2 + v_y^2 q_y^2 = m_2^2.$$

To the lowest order of $q_{x,y}$, $\epsilon(q_x^2, q_y^2)$ can be expressed as

$$\epsilon(q_x^2, q_y^2) = \frac{q_x^2}{2M_x} + \frac{q_y^2}{2M_y}.$$ (3.9)

Therefore, the above equations represent three ellipses centered at $q = 0$, denoted by $E_{1\pm}$ and $E_2$, with major axes

$$a_{1x+}^2 = 2M_x (\Delta + m_1),$$
$$a_{1y+}^2 = 2M_y (\Delta + m_1),$$
$$a_{1x-}^2 = 2M_x (\Delta - m_1),$$
$$a_{1y-}^2 = 2M_y (\Delta - m_1),$$
$$a_{2x}^2 = \frac{m_2^2}{v_x^2},$$
$$a_{2y}^2 = \frac{m_2^2}{v_y^2}.$$ (3.10)

The band crossing appears at the intersection either between $E_{1+}$ and $E_2$ or between $E_{1-}$ and $E_2$. As $m_{1,2}$ increases, $E_{1-}$ shrinks while $E_2$ grows in size, and so are guaranteed to cross each other. On the other hand, since $E_{1+}$ also expands as $m_1$ increases, it may or may not cross with $E_2$. This simple $k \cdot p$ model is only valid near $Y$ and hence is unable to reproduce the whole evolution
process of the Weyl points observed in the numerics. It can, however, satisfactorily reproduce the creation and the annihilation of the first pair of Weyl points as \(c\) decreases.

As \(m_{1,2}\) increases from zero, \(E_2\) grows from the origin, \(E_{1+}\) increases and \(E_{1-}\) decreases in size (Figure 3.16a). The pair creation at \(\Gamma Y\) happens when (Figure 3.16b)

\[
a_{2y} = a_{1y-}.
\]

(3.11)

As \(m_{1,2}\) further increases, the four crossing points between \(E_{1-}\) and \(E_2\) traces an ellipse around \(Y\) (Figure 3.16c), until the pair annihilation at \(YR\) happens when (Figure 3.16c)

\[
a_{2x} = a_{1x-}.
\]

(3.12)

Away from the \(k_z=0\) plane, one can introduce an interlayer coupling term that anti-commutes

\[
H_z = v_z \sigma_x \tau_y q_z,
\]

(3.13)

with which the dispersion becomes fully gapped

\[
E(q_x, q_y, q_z) = \sqrt{\left(\epsilon(q_x^2, q_y^2) - \Delta \pm m_1\right)^2 + \left(\sqrt{v_x^2 q_x^2 + v_y^2 q_y^2} \pm m_2\right)^2} + v_z^2 k_z^2.
\]

(3.14)

This ensures that all band crossings (Weyl points) are locked on the \(k_z=0\)-plane.

### 3.2.4 Experimental Synthesis and Mechanical Exfoliation

For the experimental synthesis and mechanical exfoliation, it is crucial to check the related materials properties such as dynamic and thermodynamic properties as well as interlayer binding energy. Figure 3.17a shows the phonon dispersion of monolayer \(\text{NbRhTe}_4\) calculated by first-principles density-functional perturbation theory, while Figure 3.18 presents the results for all four materials. Negative frequency is absent in all four monolayer cases, suggesting excellent dynamic stability upon small structural perturbation. Furthermore, we calculated the total energy difference
between the fully relaxed 1H and 1T’ structures. The results shown in Figure 3.17b reveals that 1T’ structure is more stable than 1H structure by 0.53, 0.36, 0.43 and 0.26 eV/formula unit (f.u.) for TaIrTe₄, TaRhTe₄, NbIrTe₄, and NbRhTe₄, respectively. It is thus very promising for direct experimental synthesis. In addition, the total energy difference between bulk T_d-MM’Te₄ and monolayer 1T’-MM’Te₄ displayed in Figure 3.17c yields the interlayer binding energy (Eb) of 25, 26, 26, and 27 meV/Å², respectively. All of them are even less than that of graphite (∼30 meV/Å² using the same PBE exchange-correlation functional and optB88-vdW correlation functional), validating
the feasibility of direct mechanical exfoliation to obtain their multi-layer or monolayer structures. Lastly, as shown in Figure 3.17c, all four 2D monolayer MM’Te₄ have sizable fundamental gaps of 23, 58, 30, and 65 meV for TaIrTe₄, TaRhTe₄, NbIrTe₄, and NbRhTe₄, respectively. In summary, the above results demonstrate that Peierls-distorted MM’Te₄ are both thermodynamically and dynamically stable and their interlayer binding energy is relatively weak, therefore particularly promising for experimental synthesis and mechanical exfoliation whose finite fundamental gap shall allow direct experimental measurements.

Among the four predicted topological materials, T₄₄-polytypical NbIrTe₄ and TaIrTe₄ were synthesized and characterized with the Pmn2₁ space group two decades ago [144, 145] and the measured electrical resistivity of bulk T₄₄-NbIrTe₄ and T₄₄-TaIrTe₄ increases with increasing temperature. These experiments provide a solid support for our theoretical results on thermodynamic and dynamic stabilities and semi-metallic electronic structure. It is therefore highly plausible that two other materials NbRhTe₄ and TaRhTe₄ may be synthesized as well. Moreover, it is worth to note that two metal elements M and M’ of MM’Te₄-based TTMCs possess different valence charges, making them distinctly different from convectional binary TMDCs and their alloys such as WTe₂ and W₁₋ₓMoₓTe₂. This ternary nature shall allow one to adjust the stoichiometric ratio of M and M’, thereby providing another viable strategy to manipulate the Fermi level. It will permit one to fine-tune the Fermi level position with respect to Weyl points, enabling a facile realization of chiral anomaly. Furthermore, different choices of metallic elements in TTMCs can significantly enlarge the discovery and design space for topological materials. Finally, similar atomic and electronic structures as well as the nontrivial 2D and 3D topologies between these ternary compounds and WTe₂ strongly suggest that they may also exhibit giant magnetoresistance, which is highly worth of further experimental exploration.

The present topological phase transition from stacked Z₂ TIs to Weyl semimetals originates from the breaking of inversion symmetry upon stacking, which is different from the scheme proposed by Burkov and Balents with broken TRS. [160] While finalizing the current manuscript, we also learned of two independent work by Koepernik et al. [161] and Kim et al. [162] on the
Figure 3.18: Phonon band structures of monolayer 1T’-MM’Te4.

prediction and experimental characterization of bulk Td-TaIrTe4 as a type-II Weyl semimetal, respectively, and their results are certainly encouraging. However, our work predicts that this new class of TTMC MM’Te4 (total four materials) can simultaneously serve as two-dimensional QSHIs in their monolayers and Weyl semimetals in their 3D bulk crystals. The Peierls-distorted monolayer 1T’ or bulk Td ternary structure may be realized in a more general formula with M=(Nb, Ta) and M’=(Ru, Os, Rh, Ir), providing additional materials and alloying/doping choices, meanwhile, the weak interlayer binding attraction allows them for direct mechanical exfoliation and vdW heterostructuring. [57, 61] Such rich topologies and wide possibilities as well as experimental flexibility may empower them as a unique configurable and tunable platform for studying topolog-
ical phase transition in these ternary vdW layers and developing novel quantum electronics with low power consumption/heat dissipation and quantum computing free of decoherence. In fact, the transition from QSHI to Weyl fermions in MoTe$_2$ and WTe$_2$ [74, 138] shall be similar to the present case of $T_d$-MM’Te$_4$ as the Peierls-distorted MoTe$_2$ and WTe$_2$ have exactly the same symmetry as MM’Te$_4$. When QSHI layers are far separated, their bulk is essentially fully gapped. As the interlayer distance reduces, the vdW interaction becomes non-negligible and minimizes the bandgap, eventually leads to the touching of initial massive Dirac cones and the close of the gap. With further reduced distance and strengthened interaction as well as the noncentrosymmetric nature, the gapless Dirac points subsequently split into Weyl nodes. Depending on the microscopic interaction, type-I and type-II Weyl fermions differ by the tilting of the Weyl nodes in the crystal momentum space. This essential picture is generally applicable to vdW-stacking induced topological phase transition from QSHI to Weyl semimetals. In conclusion, the present findings extend the topological materials study to a new paradigm with a much larger search and design space. It will open up a variety of unprecedented research opportunities for not only two-dimensional layered materials research, [60, 62, 77] but also the rapidly growing quest for stable topological materials that are easy for experimental synthesis and characterization as well as topological device implementation.
4. OPTICAL SECOND HARMONIC GENERATION *

Nonlinear optical properties of materials such as second and higher order harmonic generation play pivotal roles in lasers, frequency conversion, electro-optic modulators and switches, etc. The strength of nonlinear optical responses highly depends on intrinsic crystal symmetry, transition dipole moments, specific optical excitation, and local environment. Using first-principles electronic structure theory, here we predict giant second harmonic generation (SHG) in recently discovered two-dimensional (2D) ferroelectric-ferroelastic multiferroics group IV monochalcogenides (i.e., GeSe, GeS, SnSe and SnS). Remarkably, the strength of SHG susceptibility in GeSe and SnSe monolayers is more than one order of magnitude higher than that in monolayer MoS$_2$, and two orders of magnitude higher than that in monolayer hexagonal BN. Their extraordinary SHG is dominated by the large residual of two opposite intraband contributions in the SHG susceptibility. More importantly, the SHG polarization anisotropy is strongly correlated with the intrinsic ferroelastic and ferroelectric orders in group IV monochalcogenide monolayers. Our present findings provide a microscopic understanding of the large SHG susceptibility in 2D group IV monochalcogenide multiferroics from first-principles theory and open up a variety of new avenues for 2D ferroelectrics, multiferroics, and nonlinear optoelectronics, for example, realizing active electrical/mechanical switching of ferroic orders in 2D multiferroics and in situ ultrafast optical characterization of local electronic structures using noncontact noninvasive optical SHG techniques.

4.1 Second Harmonic Generation in Two-Dimensional Materials

Materials under intense optical excitation may generate large nonlinear optical responses, depending on the intrinsic crystalline symmetry, microscopic transition dipole matrix, and specific frequency and orientation of optical field applied. In centrosymmetric crystals all even order elec-

tric susceptibility tensors vanish, as both electric field and electric polarization are polar vectors with odd parity. Noncentrosymmetric materials with large second-order electric susceptibilities such as optical second harmonic generation (SHG), linear electro-optic or Pockels effect, optical rectification, and sum frequency generation are highly valuable for a variety of applications such as lasers, frequency conversion, electro-optic modulators and switches, etc. Large SHG and third harmonic generation were recently discovered in a number of two-dimensional (2D) materials, including monolayer/multilayer MoS$_2$, [163–167] MoSe$_2$, [168–170] WS$_2$, [171–173] WSe$_2$, [171, 174] hexagonal BN (h-BN), [164, 175] GaSe [176] and InSe. [177] All of them belong to the noncentrosymmetric point group D$_{3h}$ (6m2) with effectively only one independent SHG susceptibility tensor element. It has also been observed along the edge of MoS$_2$. [178] The SHG responses can be significantly enhanced by electric control [179] and resonant excitonic excitation. [180] Despite their monolayer or few layer nature, the effective nonlinear SHG susceptibility of h-BN is close to the well-developed lithium niobate nonlinear crystals widely used in integrated optical waveguide, and in MoS$_2$ monolayer it is even one order magnitude higher than that of h-BN and lithium niobate. These discoveries have enkindled further investigations of 2D nonlinear optical materials that are free of phase-matching bottleneck and promising for applications in 2D nonlinear optics, e.g. ultrathin nonlinear optical devices and spectroscopies. [179]

Very recently, group IV monochalcogenides (i.e., GeSe, GeS, SnSe and SnS, denoted by MX with M=Ge, Sn and X=Se, S) were predicted to be 2D ferroelastic-ferroelectric multiferroics [90, 181–183] with giant spontaneous electric polarization [90, 93, 181] and spontaneous lattice strain as well as sizable piezoelectricity. [70, 181] Ferroelectricity was experimentally observed by Chang et al. [105] in an MXs cousin, SnTe monolayer, with a similar type of spontaneous lattice distortion and ferroelectric polarization. Moreover, these 2D multiferroic materials possess strongly coupled structure-optical properties owing to their correlated noncentrosymmetry and ferroelastic order. [89, 181, 184–186] The presence of ferroelectricity necessitates the breaking of inversion symmetry, therefore in principle these 2D multiferroics can have finite second-order nonlinear electric susceptibilities, while specific nontrivial elements depend on the underlying
point group symmetry. These MX monolayers belong to the noncentrosymmetric point group $C_{2v}$ (mm2), and have up to five independent SHG susceptibility tensor elements. Quite interestingly, their bulk holds a centrosymmetric point group $D_{2h}$ (mmm), hence has vanishing second-order susceptibilities. In fact, MX with odd number of layers can have finite SHG response, conversely MX with even number of layers have vanishing SHG response like their infinite bulk. Nevertheless, neither theoretical nor experimental SHG has been reported for these 2D multiferroic materials.

Here, using first-principles electronic structure theory we show that these group IV monochalcogenides possess giant SHG susceptibility, much higher than that in MoS$_2$ and h-BN monolayers. The large SHG polarization anisotropy is pertinent to the ferroelastic order and noncentrosymmetric point group, and their extraordinary SHG responses are dominated by the large residual of two opposite intraband contributions in the SHG susceptibility governed by interband and intraband Berry connections. Remarkably, the SHG susceptibility in GeSe and SnSe monolayer is one order of magnitude higher than that in MoS$_2$ monolayer, and two orders of magnitude higher than that in h-BN monolayer. In contrast, MoS$_2$ and h-BN monolayers are largely determined by a single intraband term different from 2D MX. The unique ferroelasticity and ferroelectricity in these 2D multiferroics allow direct mechanical switching of lattice orientation and electrical switching of electrical polarization, accompanied with instantaneous direction and phase switching in the SHG susceptibility tensor, respectively. Our present findings not only predict colossal SHG nonlinear optical responses in group IV monochalcogenide monolayers and provide a microscopic understanding from first-principles theory, but also open up a variety of new avenues for 2D ferroelectrics, multiferroics and nonlinear optoelectronics, e.g. achieving active control of their ferroic orders with electrical and/or mechanical stimuli and in-situ ultrafast optical sensing of local electronic structures and domain evolution in these 2D multiferroics with noncontact noninvasive linear and nonlinear optical spectroscopy/imaging.

### 4.2 First-Principles Calculation of Second Harmonic Generation

Ground-state crystal structures of group IV monochalcogenide monolayers were calculated by first-principles density-functional theory (DFT) [87] implemented in the Vienna Ab initio Simula-
tion Package [147] with a plane-wave basis and the projector-augmented wave method. [148] Here we used the Perdew-Berke-Ernzerhof (PBE) form [187] of exchange-correlation functional within the generalized-gradient approximation (GGA) [188] and a Monkhorst-Pack k-point sampling for the Brillouin zone (BZ) integration. An energy cutoff of 400 eV for the plane-wave basis and a Monkhorst-Pack k-point sampling of $10 \times 10 \times 1$ were applied. Dynamical stability of the crystal structures was confirmed by phonon dispersion from density-functional perturbation theory calculations. [152] A separate stand-alone package has been developed in the group to compute the second-order SHG susceptibility tensor [31, 189, 190] interfaced with first-principles packages. A scissor operator was applied to account for the underestimated band gap in the DFT-GGA calculations. The Kramers-Kronig relation of the calculated SHG susceptibility tensors was verified, satisfying the causality condition. Benchmark calculations were performed to test the convergence with respect to number of bands, k-point sampling, etc. A dense k-point sampling of $72 \times 72 \times 1$, 40 electronic bands, and total 1000 frequency grids between the energy range of [-6 eV, 6 eV] are enough for achieving converged SHG susceptibility tensors, except that 2000 frequency grids between [-12 eV, 12 eV] were used for h-BN due to large optical gap. Only positive frequency range is presented because of the reality condition. The fundamental frequency $\omega$ in the denominator of susceptibility tensor carries a small imaginary smearing factor $\delta : \omega \rightarrow \omega + i\delta$, and $\delta = 0.05$ eV was used in this work. All symmetry-allowed and symmetry-disallowed SHG tensor elements have been extensively checked for noncentrosymmetric point groups $C_{2v}$ (mm2) and $D_{3h}$ ($\overline{6}$m2). That is, the 23 SHG susceptibility elements for $D_{3h}$ and 20 elements for $C_{2v}$ are zero at all frequencies.

Finally, as the thickness of 2D materials is not well-defined, sheet SHG susceptibility tensors (with the unit of pm²/V) are reported in this work. This is because every term of $\chi^{\text{SHG}}(-2\omega; \omega, \omega)$ requires an integration in the first BZ. The SHG susceptibility tensor can be calculated using the formulation developed by Sipe et al. in the independent particle approximation:

$$\chi^{abc}_{\text{total}} (-2\omega; \omega, \omega) = \chi^{abc}_{\text{inter}} (-2\omega; \omega, \omega) + \chi^{abc}_{\text{intra}} (-2\omega; \omega, \omega) + \chi^{abc}_{\text{mod}} (-2\omega; \omega, \omega),$$

(4.1)
containing an interband contribution, a modification on the interband linear susceptibility due to intraband motion, and a modulation on the intraband linear susceptibility by the interband motion.

More explicitly,

\[
\chi_{\text{intr}}^{abc}(-2\omega; \omega, \omega) \equiv \chi_i(\omega) + \chi_i^v(2\omega) + \chi_i^r(2\omega)
\]

\[
\chi_{\text{inter}}^{abc}(-2\omega; \omega, \omega) \equiv \chi_e(\omega) + \chi_e(2\omega), \quad \text{and}
\]

\[
\chi_{\text{mod}}^{abc}(-2\omega; \omega, \omega) \equiv \chi_m(\omega),
\]

where

\[
\chi_e(\omega) = \frac{e^3}{\hbar^2} \int \frac{dk}{4\pi^3} \left( \sum_{nm} r_{nm} \{ r_{ml} r_n \} \omega_{mn} \left[ \frac{f_{ml}}{\omega_{ml} - \omega} + \frac{f_{ln}}{\omega_{ln} - \omega} \right] \right),
\]

\[
\chi_i(\omega) = \frac{e^3}{\hbar^2} \int \frac{dk}{4\pi^3} \left( \sum_{nm} r_{nm} \{ r_{ml} r_n \} \omega_{mn} \left[ \frac{f_{ml}}{\omega_{ml} - \omega} + \frac{f_{ln}}{\omega_{ln} - \omega} \right] \right),
\]

\[
\chi_v^i(2\omega) = \frac{e^3}{\hbar^2} \int \frac{dk}{4\pi^3} \left( -8i \sum_{nm} r_{nm} \{ r_{ml} r_n \} \omega_{mn} \omega_{mn} \left[ \frac{f_{mn}}{\omega_{mn} - 2\omega} \right] \right),
\]

\[
\chi_r^i(2\omega) = \frac{e^3}{\hbar^2} \int \frac{dk}{4\pi^3} \left( 2 \sum_{nm} r_{nm} \{ r_{ml} r_n \} \omega_{mn} \omega_{mn} \left[ \frac{f_{mn}}{\omega_{mn} - 2\omega} \right] \right),
\]

and

\[
\chi_m(\omega) = \frac{e^3}{2\hbar^2} \int \frac{dk}{4\pi^3} \left( \sum_{nm} \omega_{mn} r_{nm} \{ r_{ml} r_n \} \omega_{mn} \omega_{mn} \left[ \frac{f_{mn}}{\omega_{mn} - \omega} \right] \right),
\]

\[
+ \int \frac{dk}{4\pi^3} \left( \sum_{nm} \omega_{mn} r_{nm} \{ r_{ml} r_n \} \omega_{mn} \omega_{mn} \left[ \frac{f_{mn}}{\omega_{mn} - \omega} \right] \right).
\]

In practice, this is achieved by the discrete sum with the dense k-point sampling in the first BZ: \( \int \frac{dk}{(2\pi)^3} \rightarrow \sum_k \frac{1}{V} \), where \( V \) is the volume of unit cell. For 2D materials, \( V \) becomes ill-defined for arbitrary simulation cell length along plane norm (\( L_z \) assuming \( z \) is parallel to plane normal). A more meaningful definition will be to use \( \sum_k \frac{1}{A} \) where \( A \) is the area of the 2D plane in the unit cell, and correspondingly the k-point sampling is on the 2D plane of the first BZ. It essentially leads to the following formula for “sheet” SHG susceptibility tensors (Figure 4.1) we reported in this work, which is similar to the sheet conductance/resistance used for thin-film materials and 2D
materials: $\chi_{\text{sheet}}^{\text{SHG}}(-2\omega; \omega, \omega) \times L_z$. Physically it means the ability of SHG response coming from a single 2D material. When comparing with other nonlinear optical bulk materials, one can define an effective $L_z$. This is often done by including the thickness of 2D material ($d_{2D}$) and the van der Waals thickness on both sides of 2D material (approximated by $\sim 3.4$ Å on each side), that is, $L_z = 3.4 \times 2 + d_{2D}$. However, it should be noted that this is certainly a rough estimation where the effect of symmetry and interlayer coupling has not taken into account.

Figure 4.1: Sheet SHG susceptibilities in GeSe, SnSe, GeS, SnS, MoS$_2$, and h-BN monolayers with their magnitude, real, and imaginary component.
4.3 Second Harmonic Generation in 2D Multiferroic Monolayer MX

Figure 4.2: Atomic configuration and electronic structures of group IV monochalcogenide MX monolayers. (a) Top and side views of 2D MX. Dashed orange lines indicate the primitive cell with spontaneous polarization along the y-direction. (b) The 2D surface plot of electronic band structure in GeSe monolayer in the first Brillouin zone: CBM (upper) and VBM (lower). Λ denotes the location of energy gap in the reciprocal unit cell, Λ(k_x, k_y) = [0, ±0.39] in the fractional coordinate unit. (c) Quasiparticle band structure of GeSe monolayer with the quasiparticle band gap, the lowest exciton, and the exciton binding energy labeled in the plot.

The crystal structure of group IV monochalcogenide monolayers consists of two atomic layers that are puckered along either x- or y- direction, as their nondistorted parent structure has four-fold rotation symmetry and four mirror planes. A representative 2D GeSe crystal structure is shown in Figure 4.2(a). Here, the z-axis is oriented along the plane normal of 2D MX. Our previous calculation indicates large spontaneous lattice strain η in MX monolayers depicted by a transformation strain based on the Green-Lagrange strain tensor. [181] Spontaneous tensile strain varies from 1% to 14% and the associated spontaneous compressive strain ranges from -1% to -7%. For the GeSe monolayer shown in Figure 4.2(a), its distortion is along the y direction, and consequently η_{yy} = 4.1% and η_{xx} = -2.7%. Such large spontaneous lattice strain has a nontrivial consequence in the electronic structure and crystal symmetry. Particularly, the puckering causes inversion symmetry breaking and results in the noncentrosymmetric polar point group C_{2v} (mm2),
consequently these MX monolayers exhibit large spontaneous polarization, i.e. 484, 357, 260, and 181 pC/m for GeS, GeSe, SnS, and SnSe, respectively. [181] In addition, the puckering also leads to highly anisotropic electronic band structure, as displayed in Figure 4.2(b) for the two bands near the Fermi level, namely conduction band minimum (CBM) and valence band maximum (VBM). The calculated DFT-GGA band gap is 1.13 eV at two k-points $\Lambda(k_x, k_y) = [0, \pm 0.39]$ in the fractional coordinate unit marked in (marked in Figure 4.2(b,c)).

As GeSe monolayer is a 2D semiconductor, the reduced dimensionality enhances electronhole Coulomb interaction and increases the density of states at the band edge, and consequently, large excitonic optical absorption is expected. [89,181,191] Figure 4.2(c) displays the quasiparticle band structure of GeSe monolayer using many-body quasiparticle approach [192] and first-principles quasiatom orbital method [153,193] with the exciton level marked by a red line. The computed exciton energy at the $\Lambda$ point is 1.26 eV, which is only 0.13 eV higher than the DFT-GGA gap. The small difference is due to the cancellation of quasiparticle energy correction to the DFT-GGA gap (~0.5 eV) and the large exciton binding energy (~ -0.4 eV). To correct the underestimated band gap obtained from DFT-GGA calculations, we have used the quasiparticle GW gap and exciton binding energy to correct the optical gap of monolayer MX, MoS$_2$, and BN. This is a reasonable remedy as the GW band structures are very similar to DFT band structure for the 2D materials investigated in this work. With the major difference taken into account by the scissor shift in optical band gap, the dipole matrix elements after the correction should be quite reasonable. Future efforts will be made to include quasiparticle GW electronic structure and exciton effect in a computation and memory efficient way. The unique electronic structures and multiferroic nature of monolayer MX, particularly the absence of inversion symmetry, suggests a possibility of large nonlinear optical responses in group IV monochalcogenide monolayers.

Under incident electric field $E(\omega)$ with fundamental frequency, the second-order nonlinear polarization $P(2\omega)$ at frequency $2\omega$ is determined by a third-rank electric susceptibility tensor $\chi : P_i(2\omega) = \chi_{ijk}(-2\omega; \omega, \omega) E_j(\omega) E_k(\omega)$ which is further correlated with the emitting SHG field. Figure 4.3(a-c) show the calculated sheet SHG susceptibility tensors in GeSe, MoS$_2$, and h-BN.
monolayers, and the calculated SHG data for GeS, SnSe, and SnS are presented in Figure 4.1.

\[ \chi^{(2)}_{yxx} = \chi^{(2)}_{xyx} = \chi^{(2)}_{xxy} = -\chi^{(2)}_{yyy} \]

\[ \omega \text{(eV)} \]

Re \( \chi^{(2)} \) (-2 \( \omega \); \( \omega \), \( \omega \))

Im \( \chi^{(2)} \) (-2 \( \omega \); \( \omega \), \( \omega \))

\(|\chi^{(2)}| \)

\[ 10^{-4}\text{pm}^2/\text{V} \]

\[ \omega \text{(eV)} \]

Figure 4.3: Sheet SHG susceptibilities in GeSe, MoS\(_2\), and h-BN monolayers with their magnitude, real, and imaginary component. (a) GeSe monolayer which has seven non-zero susceptibility tensor elements with five independent ones due to its point group C\(_{2v}\) (mm\(_2\)). (b) MoS\(_2\) monolayer which has four SHG elements with only one independent element due to its point group D\(_{3h}\) (\( \bar{6}m2 \)). (c) h-BN monolayer with only one independent element due to the same point group. Black dots indicate the experimental values. [165]

Both MoS\(_2\) and BN belong to the D\(_{3h}\) (\( \bar{6}m2 \)) point group with effectively only one independent nontrivial SHG susceptibility tensor element that satisfies \( \chi^{(2)}_{yxx} = \chi^{(2)}_{xyx} = \chi^{(2)}_{xxy} = -\chi^{(2)}_{yyy} \). Group IV monochalcogenides belong to the point group C\(_{2v}\) (mm\(_2\)) as mentioned earlier, thus they have five independent SHG susceptibility tensor elements: \( \chi^{(2)}_{yxx}, \chi^{(2)}_{yyy}, \chi^{(2)}_{yzz}, \chi^{(2)}_{xyx} = \chi^{(2)}_{xxy}, \chi^{(2)}_{yzz} = \chi^{(2)}_{zzz} \). The magnitude of the calculated sheet SHG susceptibility tensor element \( \chi^{(2)}_{yxx} \) is \( 5.16 \times 10^6 \) pm\(^2\)/V at 0.66 eV for monolayer GeSe, \( 3.02 \times 10^5 \) pm\(^2\)/V at 1.67 eV for monolayer MoS\(_2\), and \( 6.38 \times 10^4 \) pm\(^2\)/V at 3.22 eV h-BN, demonstrating much higher SHG response in GeSe than that in MoS\(_2\) and h-BN. Moreover, GeSe has another independent SHG tensor element \( \chi^{(2)}_{xyx} = \chi^{(2)}_{xxy} \) with a substantial magnitude.
of more than $4 \times 10^6 \text{ pm}^2/\text{V}$, and all the other three independent elements are much smaller. The computed frequency dependent SHG susceptibility satisfies the Kramers-Kronig relation as shown in Figure 4.4. It is worth noticing that, despite the fact that the energy gaps are corrected with respect to exciton energies, the computed SHG response is sensitive to the smearing factor, and the independent-particle approximation may also affect the accuracy. Consequently, the absolute magnitude may differ from experiment. Indeed, the calculated SHG response is higher than the measured SHG values MoS$_2$ shown as dots in Figure 4.3(b). Nonetheless, our results on MoS$_2$ and h-BN monolayers are in good agreement with recent works within the independent-particle approximation. [194–196] Meanwhile, the same smearing factor was used for all calculations of the six 2D materials studied to keep them on an equal footing. Very interestingly, the fundamental frequency for the first large peak of SHG susceptibility $\chi^{(2)}_{yxx}$ covers a wide range of optical spectrum, i.e. 0.66, 1.18, 0.56, and 1.00 eV for GeSe, GeS, SnSe and SnS, respectively.

Figure 4.4: Benchmark calculation of the Kramers-Kronig relation. Red and blue lines are the computed imaginary and real component of susceptibility tensor element $\chi^{(2)}_{yxx}$ of GeSe monolayer, respectively. Green dots show the real part evaluated using the imaginary part based on the Kramers-Kronig relation, in excellent agreement with the directly-calculated real component (blue lines).
Nonlinear SHG response can be characterized by shedding linearly-polarized laser beam onto the materials and measuring different polarization component of the outgoing SHG response through an analyzer. By inspecting its angular dependence, e.g. by rotating the samples, their crystallographic orientation and SHG polarization anisotropy and intensity can be determined. The SHG response in an arbitrarily oriented sample stems from all symmetry-allowed SHG nonlinear susceptibility tensor elements taking into account the relative orientation between the incident beam and the crystal lattice via proper Euler angles. Here, we consider a normal incidence geometry for the six 2D materials. As a result, the angular dependent SHG susceptibilities for the point group $C_{2v}$ (mm2) and $D_{3h}$ ($\bar{6}m2$) are given by

$$\chi^{(2)}_\parallel (\theta; C_{2v}) = \left(\chi^{(2)}_{xyx} + \chi^{(2)}_{yxx}\right) \sin \theta \cos^2 \theta + \chi^{(2)}_{yyy} \sin^3 \theta,$$

$$\chi^{(2)}_\perp (\theta; C_{2v}) = \chi^{(2)}_{yxx} \cos^3 \theta + \left(\chi^{(2)}_{yyy} - \chi^{(2)}_{xyx}\right) \cos \theta \sin^2 \theta,$$

(4.9)

$$\chi^{(2)}_\parallel (\theta; D_{3h}) = -\chi^{(2)}_{yxx} \sin 3\theta,$$

$$\chi^{(2)}_\perp (\theta; D_{3h}) = \chi^{(2)}_{yxx} \cos 3\theta,$$

where “∥” (“⊥”) stands for the polarization components of the SHG response parallel (perpendicular) to the polarization of incident electric field $E(\omega)$ at fundamental frequency $\omega$. “θ” is the rotation angle between $E(\omega)$ and the crystal lattice (e.g. the $x$ axis here). For GeSe monolayer,

$$\chi^{(2)}_{yxx} = (18.2 - 515.6 \cdot i) \times 10^4 \text{pm}^2 \text{V}^{-1}$$

$$\chi^{(2)}_{xyx} = \chi^{(2)}_{xxy} = (-232.9 - 233.5 \cdot i) \times 10^4 \text{pm}^2 \text{V}^{-1}$$

(4.10)

$$\chi^{(2)}_{yyy} = (9.0 - 11.2 \cdot i) \times 10^4 \text{pm}^2 \text{V}^{-1}$$

$$\chi^{(2)}_{yzz} = \chi^{(2)}_{zyz} = (-3.6 + 5.3 \cdot i) \times 10^4 \text{pm}^2 \text{V}^{-1}$$

$$\chi^{(2)}_{zzy} = \chi^{(2)}_{zyz} = (-1.1 + 0.1 \cdot i) \times 10^4 \text{pm}^2 \text{V}^{-1}$$
and for $\omega = 0.66$ eV, for MoS$_2$ monolayer,

$$\chi^{(2)}_{yxx} = \chi^{(2)}_{xyy} = \chi^{(2)}_{xxy} = -\chi^{(2)}_{yyy} = (\begin{pmatrix} -1.7 - 30.2 \cdot i \end{pmatrix}) \times 10^4 \text{pm}^2 / \text{V}$$

(4.11)

for $\omega=1.67$ eV, for h-BN monolayer,

$$\chi^{(2)}_{yxx} = \chi^{(2)}_{xyy} = \chi^{(2)}_{xxy} = -\chi^{(2)}_{yyy} = (\begin{pmatrix} -2.8 - 5.7 \cdot i \end{pmatrix}) \times 10^4 \text{pm}^2 / \text{V}.$$  

(4.12)

The SHG intensity is proportional to $|\chi^{(2)}(\theta)|^2$ and $|\chi^{(2)}_\perp(\theta)|^2$, hence their corresponding polar plots are presented in Figure 4.5(a-c) together with the corresponding frequencies and maximum values. The red and blue lines are $|\chi^{(2)}(\theta)|^2$ and $|\chi^{(2)}_\perp(\theta)|^2$, respectively.

Figure 4.5: Polarization anisotropy of sheet SHG susceptibilities $|\chi^{(2)}_\parallel(\theta)|^2$ and $|\chi^{(2)}_\perp(\theta)|^2$ in (a) GeSe, (b) MoS$_2$, and (c) BN marked by red and blue line, respectively. “$\parallel$” (“$\perp$”) stands for the polarization components of the SHG response parallel (perpendicular) to the polarization $E(\omega)$ of the incident electric field. “$\theta$” is the rotation angle between $E(\omega)$ and the crystal lattice (i.e. the $x$ axis here for all three 2D materials). Their maximum values are indicated by the green circles in the polar plot and listed at the bottom.
The results clearly show that the GeSe monolayer exhibits highly-polarized colossal SHG response with its maximum value located on the perpendicular component at $\theta=0$ (corresponding to $\chi_{yxz}^{(2)}$, which is also evident in the formula above. The polar plots of MoS$_2$ monolayer in Figure 3(b) precisely reflect their D$_{3h}$ symmetry, in excellent agreement with several recent experiments. Since SHG susceptibility varies with fundamental frequency, the SHG polarization anisotropy will change accordingly. The giant SHG susceptibility and polarization anisotropy in MX monolayers can be observed by noncontact noninvasive SHG spectroscopy with lower pump beam intensity, thereby achieving accurate determination of their 2D crystal orientation. In general, the total SHG susceptibility $\chi_{abc}^{\text{total}}(-2\omega; \omega, \omega)$ contains an interband contribution ($\chi_{\text{inter}}^{abc}$), a modification by intraband motion ($\chi_{\text{intra}}^{abc}$), and a modulation by interband motion ($\chi_{\text{mod}}^{abc}$), [31, 189] that is, $\chi_{\text{total}}^{abc}(-2\omega; \omega, \omega) = \chi_{\text{inter}}^{abc} + \chi_{\text{intra}}^{abc} + \chi_{\text{mod}}^{abc}$. These three contributions can be further classified into six terms, according to the location of the poles (either at $\omega$ or at $2\omega$): $\chi_{\text{intra}}^{abc} \equiv \chi_i^{(\omega)} + \chi_v^{(2\omega)} + \chi_r^{(2\omega)}$, $\chi_{\text{inter}}^{abc} \equiv \chi_e^{(\omega)} + \chi_{\text{intra}}^{abc}$, and $\chi_{\text{mod}}^{abc} \equiv \chi_m^{(\omega)}$. Detailed expressions of each term are given above. Figure 4.6(a,b) present the frequency dependent SHG susceptibility $\chi_{yxz}^{(2)}$ of GeSe monolayer and the contributions from the six terms with imaginary and real part, respectively. Among all the six terms, $\chi_v^{(2\omega)}$ and $\chi_r^{(2\omega)}$ are two leading ones responsible for the giant SHG susceptibility in the GeSe monolayer governed by interband and intraband Berry connections. The explicit expressions are given below, which were originally derived by Sipe et al. using the generalized position operator for periodic crystals [32, 36]

$$\chi_i^{v}(2\omega) = \frac{e^3}{\hbar^2} \int \frac{dk}{4\pi^3} \left( -8i \sum_{nm} \frac{r_{nm}^{a} \left( \{ v_{nm}^{b} r_{nm}^{c} \} - \{ r_{nm}^{b} v_{nm}^{c} \} \right)}{\omega_{mn}^2} \left[ \frac{f_{nm}}{\omega_{mn} - 2\omega} \right] \right),$$

$$\chi_i^{r}(2\omega) = \frac{e^3}{\hbar^2} \int \frac{dk}{4\pi^3} \left( -2i \sum_{nlm} \frac{r_{nm}^{a} \left( \{ v_{nlm}^{b} v_{nlm}^{c} \} - \{ v_{nlm}^{b} v_{nlm}^{c} \} \right)}{\omega_{mn}^2} \left[ \frac{f_{nm}}{\omega_{mn} - 2\omega} \right] \right).$$

(4.13)

The superscripts "a, b, c" indicate the Cartesian directions. The curly brackets indicate the intrinsic permutation symmetry $\chi_{abc}^{(2\omega, \omega, \omega)} = \chi_{acb}^{(2\omega, \omega, \omega)}$. Finally, $f_{nm} \equiv f_n - f_m$, where $f_n$ is occupation number of state $n$ according to the Fermi-Dirac distribution. The above two equations clearly show that $\chi_i^{v}(2\omega)$ represents the susceptibility modified by intraband motion (evident by
\(v_{mm}^b\) and \(v_{nn}^c\), where \(\chi_r^{\nu}(2\omega)\) contains the susceptibility modified by interband motion (evident by \(v_{ml}^b\) and \(v_{ln}^c\)). As shown in Figure 4.6(a-b), these two terms have large but opposite contribution. However, their sum \(\chi_i(2\omega) \equiv \chi_v^{\nu}(2\omega) + \chi_r^{\nu}(2\omega)\) remains very large, which accounts for the major part of total SHG susceptibility \(\chi_y^{(2)}\) compared to other terms including interband \(\chi_e(2\omega)\), interband \(\chi_e(\omega)\), intraband \(\chi_i(\omega)\) etc.

![Figure 4.6: Frequency-dependent (a) real and (b) imaginary part of the interband \((\chi_e(\omega), \chi_e(2\omega))\), intraband \((\chi_i(\omega), \chi_i(2\omega))\), and modulation \((\chi_m(\omega))\) contributions to the sheet SHG susceptibility \(\chi_y^{(2)}\). “\(\omega\)” and “\(2\omega\)” denote the poles of the specific terms. The blue and red dots indicate two major \(2\omega\) terms from the intraband contribution, \(\chi_v^{\nu}(2\omega)\) and \(\chi_r^{\nu}(2\omega)\). Total real and imaginary part of the SHG susceptibility \(\chi_y^{(2)}\) are marked by black line.](image-url)
To further understand the origin of the extraordinary SHG susceptibility in MX monolayers, we examine the detailed distribution of the total susceptibility $\chi^{(2)}_{yxx}$ (Figure 4.7(a,b)) and the two dominant SHG susceptibility terms, $\chi_i^v(2\omega)$ (Figure 4.7(c,d)) and $\chi_i^r(2\omega)$ (Figure 4.7(e,f)) in the first BZ. The total susceptibility $\chi^{(2)}_{yxx}$ is unambiguously concentrated around two valleys where the bandgap is located. One of the two major terms, $\chi_i^v(2\omega)$, is distributed at four spots in the first BZ close to but a bit deviated from the two valleys (Figure 4.7(c,d)), and they are largely governed by VBM and CBM near the Fermi level. The $yxx$ component of $\chi_i^v(2\omega)$ can be rearranged as follows:

$$\chi_{i,yxx}^v(2\omega) = \frac{e^3}{h^2} \int \frac{dk}{4\pi^2} \sum_{nm} 8f_{nm} \left[ \frac{-ir_{nm}^y r_{nm}^x}{\omega_{nm}^2 (\omega_{nm} - 2\omega)} \right] \Delta_{mn}^x,$$

where $\Delta_{mn} = v_{mn} - v_{nm}$ is the velocity difference between states $m$ and $n$. Figure 4.7(g) and Figure 4.7(h,i) present the vector field of $\Delta_{mn}$ and the imaginary and real part of $\left[ \frac{-ir_{nm}^y r_{nm}^x}{\omega_{nm}^2 (\omega_{nm} - 2\omega)} \right]$, respectively. The velocity difference plot shows that close to the $\Lambda$ points, $\Delta_{mn}^x(\delta k_x, \Lambda_y) = -\Delta_{mn}^x(-\delta k_x, \Lambda_y)$, and the similar behavior can be seen for the other term owing to the mirror symmetry with a (001) mirror plane. Therefore, the product of the two vector fields for specific $yxx$ component leads to the result shown in Figure 4.7(c,d). Even though the numerator $r_{nm}^y r_{nm}^x$ is only constrained by the symmetry and strength of interband Berry connection, the denominator $\omega_{mn}^2 (\omega_{mn} - 2\omega)$ becomes relatively much larger away from the bandgap at $\Lambda$. Furthermore, $\Delta_{mn}$ vanishes at two $\Lambda$ points as the velocity of both VBM and CBM vanishes. All these factors combined give rise to four spots concentrated around two valleys in the first BZ. For $\chi_i^r(2\omega)$, a detailed analysis indicates that it primarily arises from $n=$VBM, $m=$CBM, and $l=$VBM-2, CBM+1, and CBM+2, and the results are shown in Figure 4.7(j,k) accordingly. Similarly, the denominator $\omega_{mn}^2 (\omega_{mn} - 2\omega)$ largely dominate the response and concentrates it around the band gap at the two $\Lambda$ points. In summary, our results identify that the giant SHG response in GeSe monolayer is controlled by two intraband contributions $\chi_i^v(2\omega)$ and $\chi_i^r(2\omega)$ located near the two valleys. Despite the opposite contributions, the residual SHG susceptibility tensor elements $\chi^{(2)}_{yxx}$ remains extraordinary due to the much larger $\chi_i^r(2\omega)$ term.

89
Figure 4.7: The k-point dependent distribution of sheet SHG susceptibility $\chi^{(2)}_{yyxx}$ and $2\omega$ intraband terms $\chi^v_i(2\omega)$ and $\chi^r_i(2\omega)$ in the first BZ of GeSe monolayer. (a, b) Real and imaginary part of $\chi^{(2)}_{yyxx}$ dominated by the contributions from two equivalent valleys at the two bandgap locations (marked by $\Lambda$). (c-f) Real and imaginary part of $2\omega$ intraband terms $\chi^v_i(2\omega)$ and $\chi^r_i(2\omega)$. (g) The distribution of velocity difference $\Delta_{mn} \equiv v_{mm} - v_{nn}$ between CBM and VBM in the first BZ. (h-i) Real and imaginary part of non velocity terms in $\chi^v_i(2\omega)$. (j, k) Real and imaginary part of $\chi^r_i(2\omega)$ from three dominating sets of bands.

It is worth to note that bulk MX belong to point group mmm and space group Pnma, hence their centrosymmetric nature forbids SHG response. The difference between monolayer and bulk is therefore solely due to the stacking induced emerging inversion symmetry of individual non-centrosymmetric layers. As a result, multilayer MX materials shall exhibit alternating even-odd behavior. We have carried out first-principles calculations for multilayer GeSe and SnSe and the results are shown in Figure 4.1, which demonstrates that monolayer and trilayer GeSe (SnSe) have very similar SHG strength and shape at low energy range, and they start to deviate more at relatively higher frequency range with additional oscillations and peaks, largely due to the weak van der Waals stacking of multiple layers.
Figure 4.8: Strong coupling among ferroelectric ($P$), ferroelastic ($\eta$) order parameters, and sheet SHG susceptibilities $\chi^{(2)}$. (a) Four different multiferroic states in the GeSe monolayer: ($-P_x, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0$), ($P_y, \eta_{yy}, \chi_{yxx} \neq 0, \chi_{xyy} = 0$), ($-P_y, \eta_{yy}, \chi_{yxx} = 0, \chi_{xyy} = 0$), and ($P_y, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0$) marked by cyan, red, magenta, and green box, respectively. (b-e) Frequency dependent sheet SHG susceptibilities $\chi_{yxx}$ (solid line) and $\chi_{xyy}$ (dashed line) of the above four different multiferroic states in the unit of $10^4$ pm$^2$/V.

Group IV monochalcogenide monolayers are 2D ferroelastic-ferroelectric multiferroics with a few intrinsic order parameters, including the ferroelectric spontaneous polarization vector ($P$), the 2nd-rank ferroelastic spontaneous strain tensor ($\eta$), and the SHG second-order nonlinear susceptibility ($\chi^{(2)}_{abc}$ ($-2\omega; \omega, \omega$)), a 3rd-rank tensor. These order parameters are intrinsically coupled with each other. As illustrated in Figure 4.8(a), there are four possible configurations upon ferroelastic and/or ferroelectric transition. Each configuration has a peculiar set of order parameters that can be switched via ferroelastic/ferroelectric phase transition. For example, the structure at the top of Figure 4.8(a) may experience ferroelastic transition to the one on the right with order parameters evolving from ($P_y, \eta_{yy}, \chi_{yxx} \neq 0, \chi_{xyy} = 0$) to ($P_x, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0$) where $\chi_{xyy}$ (right) = $\chi_{yxx}$ (top). Their SHG susceptibility tensor will change accordingly, displayed in Figure 4.8(c) and Figure 4.8(e). A subsequent ferroelectric transition from the right to the left will change it from ($P_y, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0$) to ($-P_x, \eta_{xx}, \chi_{yxx} = 0, \chi_{xyy} \neq 0$), and its corresponding SHG susceptibility is shown in Figure 4.8(b). The switching of all the above order parameters is simply due to its inherent symmetry including lattice, sites, and electric polarization. Very importantly, our results demonstrate that the SHG susceptibility in four ferroelectric-ferroelastic multiferroic states
are distinctly different from each other, and hence the SHG intensity and phase information can be adopted as a unique noncontact measure for accessing multiferroic states. Although linear optical absorption may be used to determine the orientation of crystal lattice based on the anisotropic optical response, it cannot distinguish different ferroelectric polarization. The great sensitivity of SHG on the inversion symmetry breaking and the possibility of obtaining the SHG phase information are therefore highly advantageous.

4.4 Conclusions

In summary, our present findings provide a theoretical prediction of extraordinary SHG nonlinear optical responses and a microscopic understanding in group IV monochalcogenide monolayers. More importantly, it will open up a variety of new opportunities for 2D ferroelectrics, multiferroics, and nonlinear optoelectronics. For example, it could be of fundamental interest to see the effect of photostriction on the first and higher-order optical responses described hereby. Moreover, as the ultimate nanometer thickness of 2D materials is much smaller than their coherent wavelength, these 2D nonlinear optical materials will not suffer from the crucial phase-matching problem generally encountered in 3D nonlinear optical crystals. The extraordinary nonlinear responses together with excellent phase-matching make 2D multiferroics particularly promising for nonlinear optoelectronic devices. Very excitingly, the unique ferroelasticity and ferroelectricity in these 2D multiferroics, as illustrated in Figure 4.8(a), allow direct mechanical switching of lattice orientation and electrical switching of electrical polarization, accompanied with instantaneous direction and sign switching of the giant SHG susceptibilities. One can envisage, for example, to achieve active control of the ferroic orders in these 2D multiferroics with electrical, optical and/or mechanical stimuli while in-situ characterizing their local electronic structures and domain evolution using noncontact noninvasive ultrafast optical SHG techniques.
5. NONLINEAR PHOTOCURRENT RESPONSES

Nonlinear optical responses to external electromagnetic field, characterized by second and higher order susceptibilities, play crucial roles in nonlinear optical devices and novel optoelectronics. Herein we demonstrate the possibility to achieve nonlinear photocurrent switching in time-reversal invariant multiferroics, namely ferroicity-driven nonlinear photocurrent switching. It is enabled by a Hall-like direct photocurrent originated from the second-order current response of time-reversal invariant systems to intense electromagnetic field whose direction can be controlled by both internal ferroic orders and external light polarization. Second-order direct photocurrent consists of two types of nonlinear responses: shift current and circular photocurrent under linearly and circularly polarized light irradiation dominated by topological shift vector and Berry curvature, respectively. We elucidate the microscopic mechanism of ferroicity-driven nonlinear photocurrent switching in a representative class of 2D multiferroic materials using group theoretical analyses and first-principles electronic structure theory. The complex interplay of crystalline, permutation, gauge, and time reversal symmetries as well as inherent causality governs the symmetry properties of shift current and circular photocurrent. Ferroicity-driven nonlinear photocurrent switching combined with rich ferroic degrees of freedom in multiferroic materials will open up a variety of new avenues for realizing tunable and configurable nonlinear optoelectronics, bulk photovoltaics, nonlinear multiferroics, etc. utilizing their ferroic orders and various nonlinear responses under external field.

5.1 Introduction to Nonlinear Photocurrents

Nonlinear light-matter interaction plays a key role in the understanding, probing, and ultimate control of light and matter at the nanoscale. In particular, materials with strong nonlinear optical responses are highly desirable for many scientific disciplines and technological applica-
tions [199, 200], e.g. ultrafast nonlinear optics [6], nonlinear biosensing and imaging [7], efficient generation of entangled photon pairs for quantum computing and quantum sensing [8, 9], and all-optical transistor and computer [10–12]. Due to the odd parity of the two polar vectors, electric dipole and electric field, even order nonlinear electrical susceptibility tensors vanish in centrosymmetric crystals [4, 201], while odd order responses are not limited by such constraint.

Among a plethora of optical responses, nonequilibrium direct current (DC) from periodically driven system under light field is of particular interest. One notable example is linear photovoltaic Hall effect which has been predicted in graphene [46]. Additionally, there exist appealing second-order nonlinear DC responses to electromagnetic field in noncentrosymmetric solids, e.g. shift current (SC) and circular photocurrent (CC). Both were recently observed in Weyl semimetals [29, 202–205]. In parallel, recent experiments suggest two-dimensional van der Waals layered materials exhibit symmetry-dependent strong nonlinear optical responses, such as second/third harmonic generation. It makes nonlinear optical spectroscopy a perfect facile tool for characterizing and quantifying 2D materials, e.g. elastic strain, crystal orientation, phase transition, crystal thickness, even/odd layer oscillation, etc. Strong nonlinear responses in 2D materials also suggest their great advantage for phase-matching free nonlinear optics [186, 198, 206].

SC, known as bulk photovoltaic current [207, 208], refers to a photoexcitation of an electron from the valence to the conduction band with a simultaneous change in the phase of wave functions. Another type of nonlinear photocurrent, CC, also known as injection current, circular photogalvanic effect (CPGE) [21, 30] and quantum nonlinear optical Hall current [27, 28], arises from the interference of wavefunctions upon photoexcitation associated with a phase difference between two linearly polarized light, which allows for phase-modulated nonlinear photocurrent with tunable magnitude and direction. For example, left-circularly polarized and right-circularly polarized light can induce opposite currents. Using semiclassical Boltzmann framework, Moore et al. [27] and Sodemann et al. [28] revealed the fundamental roots of nonlinear SC and CC in Berry curvature induced anomalous velocity of metallic materials. This framework includes intraband process as the product of Berry phase and the gradient of Fermi-Dirac distribution function, equivalent to
Berry curvature dipole [28]. The nonlinear response in this case only considers the intraband process in metallic systems at the low frequency regime. In intrinsic semiconductors or insulators, the gradient of the Fermi-Dirac distribution and thus the Drude-like SC/CC response vanish. However, nonlinear photogalvanic current persists in non-centrosymmetric semiconductors due to nonlinear interband process which is absent in the above intraband model.

Herein we demonstrate the possibility to achieve nonlinear photocurrent switching in time-reversal invariant multiferroics, namely ferroicity-driven nonlinear photocurrent switching. More specifically, a Hall-like second-order direct photocurrent can be generated in time-reversal invariant systems upon electromagnetic field whose direction can be facilely controlled by both internal ferroic orders and external light polarization and chirality. We provide a microscopic picture based on first-principles theory and group theoretical analysis of crystalline symmetry, time reversal symmetry, permutation symmetry, gauge symmetry, and inherent causality. To illustrate the underlying mechanisms, we take monolayer group IV monochalcogenides (MX with M=Ge, Sn and X= S, Se, Te) as an example a ferroelectric-ferroelastic multiferroics [105, 181]. Experimentally, Chang et al. demonstrated that one of the MX compounds, atomic-thick SnTe, exhibits robust in-plane ferroelectricity at room temperature [105], and the application of voltage pulse between the scanning tunneling microscope tip and monolayer SnTe can manipulate the ferroelectric polarization through domain wall motion. Our previous theoretical study also demonstrated a wide range of kinetic barriers for coherent ferroic transition and domain wall energies [181]. These experimental and theoretical studies suggested that it is possible to realize ferroelectric/ferroelastic switching in MX monolayers. In this work, using first-principles calculations and group theoretical analyses we show that 2D MX monolayers exhibit large SC and CC responses that are dominated by topological quantities shift vector and Berry curvature, respectively. More importantly, switching ferroelastic order flips the direction of nonlinear SC and CC photocurrent by ±90°, while switching ferroelectric orders flips both nonlinear photocurrents by 180°. In addition, changing left/right circular polarization of light will induce 180° flip in CC. The microscopic understanding of nonlinear photocurrent switching from first-principles theory, together with very recent discoveries of 2D
ferroics/multiferroics as well as ferroelectric inorganic perovskites and hybrid organic-inorganic perovskites, will open a variety of new avenues for tunable and configurable nonlinear optoelectronics, bulk photovoltaics, and nonlinear multiferroics, etc.

5.2 Theory of Shift Current and Circular Current

Ferroicity-driven nonlinear photocurrent switching originates from the second-order photoinduced direct current density $\langle J_{\text{DC}} \rangle^{(2)}$. Unlike linear photocurrent, the direction of $\langle J_{\text{DC}} \rangle^{(2)}$ depends on intrinsic ferroic orders which is the key to ferroicity-driven nonlinear photocurrent switching. $\langle J_{\text{DC}} \rangle^{(2)}$ consists of two types of nonlinear photocurrents, SC and CC, which reflect the polarization change upon photoexcitation per unit volume,

$$\langle J_{\text{DC}} \rangle^{(2)} = \langle J_{\text{SC}} \rangle^{(2)} + \langle J_{\text{CC}} \rangle^{(2)},$$ (5.1)

where

$$\langle J_{\text{SC}} \rangle^{(2)} = 2\sigma_{2}^{abc}(0; \omega, -\omega) E^{b}(\omega) E^{c}(-\omega),$$ (5.2)

and

$$\frac{\langle dJ_{\text{CC}}^{a} \rangle^{(2)}}{dt} = -2 \text{Im} \eta_{2}^{abc}(0; \omega, -\omega) |E^{b}(\omega)||E^{c}(-\omega)| \sin(\phi^{b} - \phi^{c}).$$ (5.3)

The electric field can be described using phasors $E(t) = E(\omega)e^{-i\omega t} + c.c.$ For linearly polarized light $E(\omega)$ is real, while for left/right-circularly polarized light $E(\omega)$ is complex, $E^{b}(\omega) = |E^{b}(\omega)|e^{i\phi^{b}}$, with $\phi^{b} - \phi^{c} = \pm \pi/2$. We denote CC by $J_{\text{CC}}^{a,\text{C}}$ and $J_{\text{CC}}^{a,\text{C}}$ for left and right-circularly polarized light, respectively, and denote SC by $J_{\text{SC}}^{a,\text{SC}}$ and $J_{\text{SC}}^{a,\text{SC}}$ for linearly $x$-/y-polarized light, respectively.

The SC susceptibility tensor $\sigma_{2}^{abc}$ can be derived from perturbation theory [30] as shown in Chapter 2. After symmetrization of Eq. (2.54) with frequency and electric field components, we obtain

$$\sigma_{2}^{abc}(0; \omega, -\omega) = \frac{i\pi e^{3}}{2\hbar^{2}} \int [dk] \sum_{nm\sigma} f_{nm} \left( r^{b}_{mn} (r^{c}_{mn})_{;k} + r^{c}_{mn} (r^{b}_{mn})_{;k} \right) \delta(\omega_{mn} - \omega),$$ (5.4)
where \( (r^b_{nm})_{km} = \frac{\partial r^b_{nm}}{\partial k^m} - i r^b_{nm} (A^a_n - A^a_m) \) is the gauge covariant derivative. \( \int [dk] = \frac{d^d k}{(2\pi)^d} \) for \( d \) dimension. The SC susceptibility tensor under linearly polarized light can be rewritten in a more elegant expression,

\[
\sigma^{ab}_2 (0; \omega, -\omega) = -\frac{\pi e^3}{2\hbar^2} \int [dk] \sum_{nm\sigma} f_{nm} R^{ab}_{nm}(k) r^b_{nm} \delta (\omega_{mn} - \omega),
\]

where \( R^{ab}_{nm}(k) = -\frac{\partial \phi_{nm}^a(k)}{\partial k^m} + A^a_n(k) - A^a_m(k) \) is shift vector and \( \phi_{nm}(k) \) is the phase factor of the Berry connection \( r^b_{nm}(k) = |r^b_{nm}(k)|e^{i\phi_{nm}(k)}. \) \( r^b_{nm}r^b_{mn} \) is the optical absorption strength. Hence, SC is determined by the product of linear photoabsorption and shift vector integrated over the Brillouin zone.

The CC susceptibility tensor \( \eta^{abc}_2 \) [30] is given by the symmetrization of Eq. (2.58)

\[
\eta^{abc}_2(0; \omega, -\omega) = -\frac{\pi e^3}{2\hbar^2} \int [dk] \sum_{nm\sigma} \Delta^a_{nm} f_{nm} [r^b_{nm}, r^c_{nm}] \delta (\omega_{mn} - \omega),
\]

where \( [r^b_{mn}, r^c_{nm}] \equiv r^b_{nm}r^c_{nm} - r^c_{mn}r^b_{mn}, \) indicating CC vanishes under linearly polarized light. It can be rewritten in a general form assuming light propagates along \( z, \)

\[
\eta^{a,z}_2(0; \omega, -\omega) = \frac{i\pi e^3}{2\hbar^2} \int [dk] \sum_{nm\sigma} f_{nm} \Delta^a_{nm} \Omega^z_{mn}(k) \delta (\omega_{mn} - \omega).
\]

\( \Omega^z_{mn}(k) \equiv i [r^x_{mn}, r^y_{nm}] = -i [r^y_{mn}, r^x_{nm}] \) is local Berry curvature between band \( m \) and \( n. \) The global one reads \( \Omega^z_m = \sum_{n=m} \Omega^z_{mn}. \) Its clear to see \( \eta^{a,z}_2 = \eta^{a,z}_2. \) The original \( \text{"bc" indices} \) in the CC susceptibility tensor are now absorbed in the index \( \text{"z"}. \)

SC and CC involve distinct physical processes. Figure 5.1 shows the corresponding microscopic picture using a two-band model. SC arises from the displacement of wave packet upon photoabsorption, while CC stems from the asymmetric transport of electrons and holes and the self-rotation of the wave packet. The latter induces orbital magnetic momentum coupled with the circularly polarized light. The intrinsic permutation symmetry of electric field leads to the permutation relation \( \chi^{abc}_2(-\omega_m - \omega_n; \omega_m, \omega_n) = \chi^{abc}_2(-\omega_m - \omega_n; \omega_n, \omega_m), \) where \( \chi^{abc}_2 \) is a general second-
order susceptibility. If time reversal symmetry is also present, $r^{b}_{nm}(-k) = r^{b}_{mn}(k)$, $R^{a,b}_{mn}(-k) = -R^{a,b}_{nm}(k)$ and $\Omega^{z}_{mn}(k) = -\Omega^{z}_{mn}(-k)$. As a result of the causality and permutation symmetry, $\eta_{2}^{abc}(0;\omega,-\omega) = \eta_{2}^{abc}(0;-\omega,\omega) = -\eta_{2}^{abc}(0;-\omega,\omega) = [\eta_{2}^{abc}(0;-\omega,\omega)]^{*}$, ensuring that $\eta_{2}^{abc}$ is purely imaginary. In contrast, $\sigma_{2}^{abc}(0;\omega,-\omega) = \sigma_{2}^{abc}(0;-\omega,\omega) = \sigma_{2}^{abc}(0;-\omega,\omega) = [\sigma_{2}^{abc}(0;-\omega,\omega)]^{*}$, suggesting that $\sigma_{2}^{abc}$ is purely real.

![Microscopic interpretation of shift current and circular current using two-band model](image.png)

Figure 5.1: Microscopic interpretation of shift current and circular current using two-band model. $\delta \langle a \rangle$ is the variation of the mean value of position operator indicating the shift of electron wave packet in real space. Photoexcitation induces the shift of the electron wave packet in real space. SC comes from the displacement of wave packet upon photoabsorption, while CC stems from the asymmetric motion of electrons and holes and the self-rotation of the wave packet. The latter induces itinerant orbital magnetic momentum coupled with the circularly polarized light. Transition rate $W$ of SC is proportional to the linear optical absorption strength under linearly polarized light at frequency $\omega$, while transition rate $F$ of CC is proportional to the local Berry curvature under circularly polarized light at frequency $\omega$. 

98
Although SC and CC have different physical meaning, they are closely related. The derivative of CC susceptibility tensor $\eta_{2, z}^{\alpha, z}$ in Eq. (5.7) can be written as the following

$$
\partial_\omega \eta_{2, z}^{\alpha, z}(0; \omega, -\omega) = \frac{i \pi e^3}{2h^2} \int [dk] \sum_{nm\sigma} f_{nm} (\partial_{k^x} \Omega_{mn}^z) \delta (\omega_{mn} - \omega),
$$

(5.8)

where the integration by parts is applied. Here $\partial_{k^x} \Omega_{mn}^z$ is the Berry curvature dipole. Furthermore,

$$
\nabla_k \times R_{mn} \cdot \hat{z} = \frac{\partial R_{mn}^{y, a}(k)}{\partial k^x} - \frac{\partial R_{mn}^{x, a}(k)}{\partial k^y} = \Omega_m^z - \Omega_n^z.
$$

(5.9)

It shows that the two topological quantities, shift vector and Berry curvature, are closely connected. For a two-band model, $\Omega_m^z - \Omega_n^z = 2\Omega_{mn}^z$, thus

$$
\partial_\omega \eta_{2, z}^{\alpha, z}(0; \omega, -\omega) = -\frac{i \pi e^3}{4h^2} \int [dk] \sum_{\sigma} \nabla_k a \left( \nabla_k \times R_{wc} \cdot \hat{z} \right) \delta (\omega_{cw} - \omega).
$$

(5.10)

Therefore, both SC and the derivative of CC with respect to frequency are related to the shift vector, hence the Berry curvature. The derivative of susceptibility $\partial_\omega \eta_{2, z}^{\alpha, z}$ will contribute to the temporal response of rectification current rate when a short laser pulse is applied [209].

### 5.3 First-Principles Atomistic and Electronic Structure Calculation Methods

Ground-state crystal structures of MX were calculated by first-principles density-functional theory implemented in the Vienna Ab initio Simulation Package (VASP) with the projector-augmented wave method and a plane-wave basis with an energy cutoff of 400 eV. We employed the generalized-gradient approximation of exchange-correlation functional in the Perdew-Burke-Ernzerhof form and a Monkhorst-Pack k-point sampling of $12 \times 12 \times 1$ for the Brillouin zone integration.

To evaluate the nonlinear photocurrent responses, we developed a nonlinear optical code interfaced with first-principles density functional theory packages (e.g. VASP and Quantum-ESPRESSO). Second-order nonlinear photocurrent responses such as SC and CC are different from linear optical absorption, as the numerical integration of generalized Berry connection has to be performed over the whole Brillouin zone and multiple band indices for 27 third-rank tensor elements at each fre-
quency. We therefore parallelized and benchmarked the code with tensor symmetrization and SOC taken into account. For the present calculations, we found that a dense 72×72×1 k-point sampling, 40 electronic bands without SOC (80 bands with SOC), and total 1000 frequency grids in an energy range of [-6 eV, 6 eV] are enough to achieve converged nonlinear photocurrent susceptibility tensors. The fundamental frequency \( \omega \) in the denominator of susceptibility tensor carries a small imaginary smearing factor \( \tau \) of 0.05 eV: \( \omega \rightarrow \omega + i\tau \). We used the Sokhotski-Plemelj theorem to perform the integrals of Dirac delta function, which reads \( \lim_{\tau \to 0} \frac{1}{\omega_{mn} \pm (\omega + i\tau)} = \mathcal{P} \frac{1}{\omega_{mn} - \omega} \mp i\pi \delta (\omega_{mn} \pm \omega) \). \( \mathcal{P} \) is the Cauchy principal value. It should be emphasized that this formula is only meaningful when it is multiplied by a function and integrated over the first Brillouin zone and the real line that includes the point \( \omega = \omega_{mn} \). This always holds true in our case. Finally, the single-particle approximation often underestimates the bandgap, and two-body and higher-order interactions may also affect the results. To check the effect of the single-particle approximation, we applied a scissor operator to correct both bandgap and matrix elements by using the optical bandgap obtained from quasiparticle GW and Bethe-Salpeter equation calculations. This approach is a reasonable remedy as the GW and DFT band structures are similar for the MX materials. Figure 5.2 shows the nonlinear photocurrent calculated with and without the scissor operator, and the main effect is the relative shift of the peak positions with slightly reduced peak intensity. This is expected as it mostly affects the denominator and delta function in the equations for SC and CC. Two-body and higher-order excitations in principle can be taken into account by incorporating many-body wavefunctions (e.g. exciton wavefunctions as a linear combination of electron-hole pair excitations) into the SC and CC tensor, which may change the SC and CC peak position and amplitude. Nonetheless, the main conclusions on the ferroicity-driven nonlinear photocurrent switching behavior from the symmetry analysis and first-principles calculations shall still hold, and future efforts shall be made to include the exciton and other high-order excitation effect.
Figure 5.2: Shift current susceptibility tensor elements ($\sigma_{yx}$, $\sigma_{yy}$), and circular photocurrent susceptibility tensor elements ($\eta_{xy}$, $\eta_{yx}$) of monolayer GeS with and without scissor operator. The scissor operator is based on the bandgap correction using quasiparticle GW gap and exciton binding energy.

5.4 First-Principles Calculations and Group Theoretical Analysis of Nonlinear Photocurrent Responses

The symmetry property of linear susceptibility and nonlinear SC and CC susceptibility are governed by point group and permutation symmetry, which correspond to direct product $\Gamma_{P} \otimes \Gamma_{E}$, $\Gamma_{jSC} \otimes \Gamma_{EE}$, $\Gamma_{jCC} \otimes \Gamma_{EE}^*$, respectively. Here we take monolayer group IV monochalcogenides (MX with M=Ge, Sn and X=Se, S) as an example which is a ferroelectric-ferroelastic multiferroics [181]. In 2D MX with C$_{2v}$ point group (see Figure 5.5A and 5.5B for crystal and electronic structure, respectively), we have $\Gamma_{P} \otimes \Gamma_{E} = 3A_1 + 2A_2 + 2B_1 + 2B_2$. Hence there are three independent nonzero components in linear susceptibilities. The permutation symmetry further separates them into symmetric and asymmetric representations [55], $\Gamma_{P} \otimes \Gamma_{E} = \Gamma^s + \Gamma^a$, where $\Gamma^s = 3A_1 + A_2 + B_1 + B_2$ and $\Gamma^a = A_2 + B_1 + B_2$. Moreover, since polarization $P$, current $J$, and electric field $E$ are all polar vectors, they share the same representation, thus
\[ \Gamma_{J_{SC}} \otimes \Gamma_{EE} = 7A_1 + 6A_2 + 7B_1 + 7B_2, \text{ and } \Gamma_{J_{SC}}^s = 5A_1 + 3A_2 + 5B_1 + 5B_2. \] As a result, there are five independent nonzero components in the SC susceptibility tensor. Figures 5.5C and 5.5D show two of the five nontrivial frequency dependent SC susceptibilities \( \sigma_{\frac{1}{2}^{yx}}(0; \omega, -\omega) \) and \( \frac{1}{2}^{yy}(0; \omega, -\omega) \) in 2D MX where linearly \( x \)- and \( y \)-polarized light are considered. The corresponding SC along \( y \) direction reads \( J_{SC}^{yy} = 2\sigma_{\frac{1}{2}^{yx}} E_x E_x \) and \( J_{SC}^{yy} = 2\sigma_{\frac{1}{2}^{yy}} E_y E_y \), respectively. The SC for monolayer GeS agrees well with the results in Refs. (21, 32). Furthermore, as shown in Figure 5.5C and 5.5D, spin-orbit coupling (SOC) only slightly affects the SC in the case of GeS because of weak spin-orbit interaction strength of Ge and S atoms. \( \sigma_{\frac{1}{2}^{yy}} \) has two peaks below frequency of 3 eV. The first peak at 2 eV (denoted by red circle in Figure 5.5C) is contributed from the k-points around the Brillouin zone center (\( \Gamma \) point). The second peak at 2.8 eV denoted by green circle is apparently very strong. It comes from the transition at the k-points around the Brillouin zone boundary (X point) in a butterfly shape as shown in Figure 5.5C. However, SOC can have significant impact in other cases such as well-known 1H-MoSe\(_2\) and WSe\(_2\), and the results for all MX and MX\(_2\) without and with SOC are shown in Figure 5.3 and Figure 5.4.

The reciprocal vector dependent contributions to SC are shown in Figure 5.5E and 5.5F for photon energy of 2.0 and 2.8 eV, respectively. They are determined by the product of SC susceptibility density, \( \text{Im}[r_{vc}^{b} r_{cv}^{b}; k\alpha] \) (Figure 5.5G and 5.5J for 2.0 and 2.8 eV, respectively), and the energy conservation law is carried by \( \delta(\omega_{cv} - \omega) \). The distribution of shift current susceptibility density can be further understood by focusing on the frequency independent terms determined by the product of dipole transition strength, \( r_{vc}^{b} r_{cv}^{b} \) (i.e. optical absorption), and shift vector \( R_{nm}^{a,b}(k) \), as shown in Figure 5.5H-5.5I and 25.5K-5.5L, for 2.0 and 2.8 eV, respectively. \( \text{Im}[r_{vc}^{b} r_{cv}^{b}] \) vanishes at the k points around the band gap, while \( \text{Im}[r_{vc}^{b} r_{cv}^{b}] \) remains finite due to the optical selection rule. \( R_{nm}^{a,b}(k) \) is a gauge invariant topological quantity which is well defined away from optical zero points, i.e. \( r_{nm}^{b}(k) \neq 0 \). Since the optical zero points have no contribution to SC, we compute the shift vector by

\[
R_{nm}^{a,b}(k) = \frac{1}{|r_{nm}^{b}|^2} \text{Im}[r_{mn}^{b} r_{nm}^{b}; k\alpha] \quad (5.11)
\]

The quantitative relationship between shift vector and polarization difference was recently re-
Figure 5.3: Shift current and circular photocurrent in monolayer MX \((C_{2v})\) and 1H-MX\(_2\) \((D_{3h})\) without spin orbit coupling. (A) Shift current susceptibility in 1H-MX\(_2\). (B) Shift current susceptibility in MX. (C) Circular photocurrent susceptibility in MX.

Figure 5.4: Shift current and circular photocurrent in monolayer MX \((C_{2v})\) and 1H-MX\(_2\) \((D_{3h})\) with spin orbit coupling. (A) Shift current susceptibility in 1H-MX\(_2\). (B) Shift current susceptibility in MX. (C) Circular photocurrent susceptibility in MX.
ported [210]. In the presence of time reversal symmetry, \( r_{nm}^b(-k) = r_{nm}^b(k), (r_{nm}^b(-k));_{k\alpha} = -(r_{nm}^b(k));_{k\alpha} \), which leads to \( P_{nm}^{a,b}(-k) = -P_{nm}^{a,b}(k) = P_{nm}^{a,b}(k) \). This is clearly confirmed in Figure 5.5I and 5.5L. The shift vector can reach as high as \( \sim 15 \) Å, much larger than its lattice constant. This is very different from electric polarization vector which is smaller than the lattice vector. Figures 5.6 and 5.7 show the distribution of the SC susceptibility tensor elements for monolayer MoSe\(_2\) in the first Brillouin zone, demonstrating that its shift vector defined in Eq. (5.11) can go beyond its lattice parameter.

Figure 5.5: Shift current (SC) and its microscopic origin in 2D ferroelastic-ferroelectric monolayer group IV monochalcogenide GeS. (A) Crystal structure of monolayer group IV monochalcogenides MX where M=(Ge, Sn) and X=(S, Se). (B) 2D electronic band structure near the Fermi level. (C) and (D) Frequency-dependent nonlinear shift current response to incoming linearly \( x \) and \( y \) polarized light, respectively. (E) and (F) Reciprocal vector \( k \) resolved SC susceptibility under linearly \( y \) polarized light at the first two peaks (2.0 and 2.8 eV). (G) and (J) \( k \)-resolved SC strength, (H) and (K) \( k \)-resolved dipole transition strength, (I) and (L) \( k \)-resolved topological shift vector of 2D GeS in 2D Brillouin zone under linearly \( x \) and \( y \) polarized light, respectively.
Figure 5.6: Microscopic distribution and frequency dependent shift photocurrent susceptibility in monolayer 1H-MoSe$_2$ with D$_{3h}$ point group. (A, E) Shift current susceptibility tensor elements $\sigma_{yx}^2$ and $\sigma_{yy}^2$ for shift current using four bands and SOC. (B-D, F-H) the distribution in k-space contributed to the susceptibility tensor element under incident light with different frequencies.

For circularly polarized light along $z$, $\mathbf{E} \times \mathbf{E}^*$ and axial axis $R_z$ share the $B_2$ representation. Therefore, $\Gamma_{\mathbf{j}^x} \otimes \Gamma_{R_z} = B_2 \otimes B_2 = A_1$, indicating there is a nonzero CC response along $x$ direction $\eta_{2}^{x,z}(0; \omega, -\omega)$ when the applied circularly polarized light is along $z$ direction, i.e. perpendicular to 2D plane. Figures 5.8A and 5.8B show two antisymmetric CC susceptibility tensor elements, $\text{Im} \eta_{2}^{xy}$ and $\text{Im} \eta_{2}^{yx}$, respectively. Unlike SC, the main response of CC spread in two peaks from 2 to 6 eV. The peak values of $\text{Im} \eta_{2}^{yx}$ in the four MX materials are about $100\sim 300 \times 10^8$ nm AV$^{-2}$ s$^{-1}$, which allows us to estimate the generated nonlinear circular current under continuous wave limit as follows. At room temperature, a typical relaxation time of the electrons in MX materials is around $\tau = \mu_e/\left(\frac{e}{m^*}\right) \sim 10^{-14}$ s (34). $\mu_e$ is the mobility and $m^*$ is the effective mass of electrons. Considering a regular laser pointer with an intensity of 1 mW/cm$^2$ and 2D MX with an effective thickness of 1 nm, the induced circular photocurrent $J_{CC}$ can reach $10\sim 30 \mu$A/cm$^2$, indicating...
Figure 5.7: Shift current and its microscopic origin in monolayer 1H-MoSe$_2$. (A) 2D electronic band structure near the Fermi level along with its crystal structure. (B) and (E) $k$-resolved shift current strength, (C) and (F) $k$-resolved dipole transition strength, (D) and (G) $k$-resolved topological shift vector of monolayer 1H-MoSe$_2$ in 2D Brillouin zone under linearly $x$ and $y$ polarized light, respectively.

The current can even be observed using low power continuous wave source (sheet photocurrent of 10−30 nm $\mu$A/cm$^2$, that is, 100−300 pA/m) [211]. SC and CC are generated simultaneously under circularly polarized light. It is possible to compare their peak amplitudes if we assume the same incident light intensity and a relaxation time of $10^{-14}$ s for GeS. CC is larger than SC, $J_{CC}/J_{SC} = \sim 5$. The second-order nonlinear photocurrent response for different incident polarized light is summarized in Table 5.1. It should be noted that there is another nonzero element $\eta_{z,x}^x$ in the CC susceptibility tensor since $\Gamma_j \otimes \Gamma_{R_x} = B_1 \otimes B_1 = A_1$, suggesting there exists a CC response along $z$ direction when the incoming circularly polarized light is along $x$ direction.

Figures 5.8C and 5.8D show the $k$ resolved CC susceptibility in monolayer GeS under circularly polarized light at two different frequencies (2.3 and 2.8 eV), demonstrating that the main response of the CC is localized around Y point. It should be noted that, for the same frequency of 2.8 eV, the SC (Figure 5.5F) and CC (Figure 5.8D) are very different from each other, as the
Table 5.1: Second order nonlinear photocurrent responses under different polarized light. ↔ and ↓ indicate horizontal \((x)\) and vertical \((y)\) linearly-polarized light, respectively. \(\bigcirc\) and \(\bigotimes\) indicate left and right circularly polarized light, respectively. \(E_0\) is the strength of the applied electric field. \(\varphi\) stands for linearly-polarized light that has a 45 degree angle with respect to \(x\) and \(y\). Shift current (SC) and circular photocurrent (CC) can coexist in 2D MX (with \(C_{2v}\) point group) upon left/right circularly polarized light. However, only SC will persist, while CC will disappear in 2D MX under linearly polarized light such as ↔, ↓, and \(\varphi\). It is important to mention that both SC and CC in monolayer MX\(_2\) (e.g. 1H-MoSe\(_2\) with \(D_{3h}\) point group) will be always zero under left/right circularly polarized light.

<table>
<thead>
<tr>
<th>Light polarization</th>
<th>SC</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>↔</td>
<td>(J_{SC}^{\leftrightarrow} = 2\sigma_{yxx} E_0^2)</td>
<td>(J_{CC}^{\leftrightarrow} = 0)</td>
</tr>
<tr>
<td>↓</td>
<td>(J_{SC}^{\downarrow} = 2\sigma_{yyy} E_0^2)</td>
<td>(J_{CC}^{\downarrow} = 0)</td>
</tr>
<tr>
<td>(\varphi) (= \frac{\leftrightarrow + \downarrow}{\sqrt{2}})</td>
<td>(J_{SC}^{\varphi} = (\sigma_{yxx} + \sigma_{yyy}) E_0^2)</td>
<td>(J_{CC}^{\varphi} = 0)</td>
</tr>
<tr>
<td>(\bigcirc) (= \frac{\leftrightarrow - \downarrow}{\sqrt{2}})</td>
<td>(J_{SC}^{\bigcirc} = (\sigma_{yxx} + \sigma_{yyy}) E_0^2)</td>
<td>(J_{CC}^{\bigcirc} = 2 \text{Im} \eta_{2y}^2 E_0^2)</td>
</tr>
<tr>
<td>(\bigotimes) (= \frac{\leftrightarrow - i \downarrow}{\sqrt{2}})</td>
<td>(J_{SC}^{\bigotimes} = (\sigma_{yxx} + \sigma_{yyy}) E_0^2)</td>
<td>(J_{CC}^{\bigotimes} = 2 \text{Im} \eta_{2y}^2 E_0^2 = -J_{CC}^{\bigcirc})</td>
</tr>
</tbody>
</table>

SC is localized around Y point. The CC susceptibility is determined by the product of group velocity difference and Berry curvature \(\Omega_{cv}^z(k)\). The susceptibility tensor relates the component of the polar vector \(J\) and the axial vector \(\mathbf{e} \times \mathbf{e}^*\). Hence, it is nonzero for the point groups that allow optical activity or gyrotropy. Here \(\mathbf{e}\) is the unit vector of light polarization. Figure 5.8E shows the group velocity difference between the highest valence band and the lowest conduction band, which confirms the time reversal symmetry \(\Delta(k) = -\Delta(-k)\) (see APPENDIX X for the energy-dependent group velocity distribution). The Berry curvature \(\Omega_{cv}^z(k)\) of GeS is shown in Figure 5.8F, which confirms \(\Omega_{cv}^z(k) = -\Omega_{cv}^z(-k)\) under time reversal symmetry and the mirror plane \((yz\)-plane) leads to opposite Berry curvature at \((\pm k^x, k^y)\). The product of these two odd functions, \(\Delta(k)\) and \(\Omega_{cv}^z(k)\), results in nonvanishing CC in 2D MX with \(C_{2v}\) point group. This is in direct constrast to 1H-MoSe\(_2\) whose CC response vanishes due to its \(D_{3h}\) point group as evident in its Berry curvature shown in Figure 5.9.

5.5 Ferroicity-Driven Nonlinear Photocurrent Switching

The above group theoretical analyses and first-principles calculations illustrate the underlying selection rule and microscopic mechanism governing nonlinear SC and CC photocurrents.
Figure 5.8: Circular photocurrent (CC) and its microscopic origin in 2D ferroelastic-ferroelectric monolayer group IV monochalcogenide GeS. (A) and (B) Two opposite CC susceptibility tensor elements induced by circularly-polarized light. (C) and (D) Evolution of reciprocal vector \( k \) resolved CC susceptibility under circularly polarized light at two different frequencies (2.3 and 2.8 eV). (E) and (F) Group velocity difference and Berry curvature between the highest valence band and the lowest conduction band. The white arrows in (E) denote the calculated group velocity difference at specific \( k \) point, and the black curves indicate the associated stream lines.

Since they are intimately related to the symmetry and topology, nonlinear SC and CC photocurrents are inherently coupled with the intrinsic ferroic orders in 2D multiferroics MX, giving rise to ferroicity-driven nonlinear photocurrent switching which we will discuss below. Let's first inspect the coupling between ferroelectric order \( (P_y) \) and nonlinear SC and CC responses. Since both CC and SC are polar vector, the sign of SC will flip upon ferroelectric polarization switch \( (P_y \rightarrow -P_y) \). Consequently, the sign of SC susceptibility tensor \( \sigma^{ybb}_2(0; \omega, -\omega) \) and CC susceptibility tensor \( \eta^{x,z}_2((0; \omega, -\omega)) \) will also flip. Thus, under the same linearly/circularly polarized light, SC and CC will change the direction by 180° upon ferroelectric polarization switch as shown in Figure 5.10, that is, \( J^{y,*}_{SC}(P_y) = -J^{y,*}_{SC}(-P_y) \), \( J^{y,\downarrow}_{SC}(P_y) = -J^{y,\downarrow}_{SC}(-P_y) \), \( J^{x,\odot}_{CC}(P_y) = -J^{x,\odot}_{CC}(-P_y) \), and \( J^{x,\odot}_{CC}(P_y) = -J^{x,\odot}_{CC}(-P_y) \). Such property can be obtained from microscopic theory by consid-
Figure 5.9: Berry curvature distribution in monolayer 1H-MoSe\(_2\).

erring the transformation rules of different matrix elements under space inversion and time reversal operation, including inter-band Berry connections \(r_{mn}(k)\), shift vector \(R_{mn}(k)\), group velocity difference \(\Delta_{mn}(k)\), and Berry curvature \(\Omega_{mn}(k)\) as listed in Table 5.2.

Next we examine the coupling between ferroelastic order (e.g. spontaneous strain \(\epsilon_{yy} > 0\) and \(\epsilon_{xx} < 0\)) and nonlinear photocurrent responses. Upon ferroelastic transition \((\epsilon_{yy} \rightarrow \epsilon_{xx} \text{ and } \epsilon_{xx} \rightarrow \epsilon_{yy})\), shift vector \(R_{nm}(k)\), Berry curvature \(\Omega_{mn}(k)\), \(r_{mn}^{b,b}\) and \(\Delta_{mn}^{a}\) will all rotate by 90°, which effectively switches the \(xy\) index. As a result, under the same linearly/circularly polarized light, nonlinear SC and CC photocurrent will change their direction by 90° upon ferroelastic transition, that is, \(J_{SC}^{y,\leftrightarrow}(\epsilon_{yy}) = J_{SC}^{x,\leftrightarrow}(\epsilon_{xx})\), \(J_{SC}^{y,\uparrow}(\epsilon_{yy}) = J_{SC}^{x,\uparrow}(\epsilon_{xx})\), \(J_{CC}^{\bigcirc}(\epsilon_{yy}) = J_{CC}^{\bigcirc}(\epsilon_{xx})\), and \(J_{CC}^{\bigcirc}(\epsilon_{yy}) = J_{CC}^{\bigcirc}(\epsilon_{xx})\). Note that in general \(J_{CC}^{\bigcirc} = -J_{CC}^{\bigcirc}\), \(J_{SC}^{y,\leftrightarrow}\), and \(J_{SC}^{y,\uparrow}\) are independent. However, \(J_{SC}^{y,\leftrightarrow} = -J_{SC}^{y,\uparrow}\) holds in group \(D_{3h}\) with a mirror plane perpendicular to \(x\) axis, e.g. 1H-MoSe\(_2\).
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Space Inversion, $\mathcal{I}$</th>
<th>Time Reversal, $\mathcal{T}$</th>
<th>Gauge dependency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{mn}(k)$</td>
<td>$-r_{mn}(-k)$</td>
<td>$r_{mn}(-k)$</td>
<td>Yes</td>
</tr>
<tr>
<td>$R_{mn}(k)$</td>
<td>$-R_{mn}(-k)$</td>
<td>$R_{mn}(-k)$</td>
<td>No</td>
</tr>
<tr>
<td>$\Omega_{mn}(k)$</td>
<td>$-\Omega_{mn}(-k)$</td>
<td>$-\Omega_{mn}(-k)$</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 5.2: Transformation of inter-band Berry connections $r_{mn}$, shift vector $R_{mn}$, group velocity difference $\Delta_{mn}$ and Berry curvature $\Omega_{mn}$ under space inversion $\mathcal{I}$ and time reversal $\mathcal{T}$ symmetry operation. $R_{mn}(k)$ is odd under $\mathcal{I}$ and even under $\mathcal{T}$ in moment space. $\Omega_{mn}(k)$ is even under $\mathcal{I}$ and odd under $\mathcal{T}$ in moment space. These transformation rules govern the coupling between ferroelectric polarization and nonlinear SC and CC photocurrent.

Since 2D MX possess both ferroelectric and ferroelastic orders, it has four multiferroic ($\pm P, \pm \epsilon$) states whose nonlinear photocurrent SC and CC are directly correlated. Here we define ferroelastic strain $+\epsilon$ for $\epsilon_{xx} < 0$ and $\epsilon_{yy} > 0$, and $-\epsilon$ for $\epsilon_{xx} > 0$ and $\epsilon_{yy} < 0$. Ferroelectric polarization $P$ could be $\pm P_x$ if $\epsilon_{xx} > 0$, or $\pm P_y$ if $\epsilon_{yy} > 0$. In contrast, linear optical susceptibility will not change with ferroelectric polarization switching because its matrix element $r_{mn}^b r_{mn}^b$ is always positive. Moreover, for a given multiferroic state $(P, \epsilon)$, nonlinear SC and CC current responses in 2D MX are bulk photocurrent response along different directions, thus SC and CC can serve as a fundamental principle for real-space mapping of ferroelectric and ferroelastic orders.
5.6 Generalization of Nonlinear Photocurrent Switching

The above group theoretical analysis and first-principles calculations can be generally applied to other ferroic materials that can host ferroicity-driven nonlinear photocurrent switching. We can extend the above analysis to all 32 crystallographic point groups. 21 are noncentrosymmetric point groups, 10 of which ($C_1, C_s, C_n, C_{nv}, n=2,3,4,6$) possess polar axis a necessary but insufficient condition for ferroelectric materials. Assuming $z$ to be the polar axis, $z$ will transform as the total symmetric representation, therefore $\Gamma_{jz} \otimes \Gamma_{Ex} \otimes \Gamma_{Ex}$ and $\Gamma_{jz} \otimes \Gamma_{Ey} \otimes \Gamma_{Ey}$ always contain the total symmetric representation, indicating that nonvanishing SC is perpendicular to electric field and thus leads to a nonlinear transverse photocurrent. Such transverse current cannot be observed from conventional linear Hall effect in time-reversal invariant materials. Furthermore, 18 of 21 noncentrosymmetric point groups including all the 10 polar classes are gyrotropic, hence they all possess nonvanishing CC. Although our focus here is on the ferroicity-driven nonlinear photocurrent switching based on interband-process induced second-order photocurrent, the similar effect can be also found in ferroelectric semimetals where static/low-frequency electric field induced nonlinear Hall current can be switched upon ferroelectric switching. However, it should be emphasized that not all ferroic materials exhibit nonlinear photocurrent switching upon ferroic transition. Similar to the present case, it depends on the underlying symmetry transformation of ferroelectric transition. More excitingly, it can be readily extended to higher-order or other types of nonlinear effects, for examples, optically induced nonlinear magnetization switching and ferroicity-driven third-order nonlinear photocurrent switching. It will offer unprecedented opportunities for nonlinear optoelectronics, nonlinear optomagnetics, and nonlinear optoelectromagnetics, namely nonlinear multiferroics, as potential routes to control polarization, charge current, magnetization, and spin current in an ultrafast, noninvasive manner without the stringent requirement of strongly-coupled lattice, polarization, and magnetization in conventional multiferroics.

In summary, using group theoretical analyses and first-principles calculations we have studied the microscopic mechanism of nonlinear photocurrent switching in time-reversal invariant multiferroics. Our results show that nonlinear photocurrent is highly sensitive to the symmetry of
materials, including point group symmetry, permutation symmetry, and time reversal symmetry. This leads to ferroicity-driven nonlinear photocurrent switching unique to multiferroics where the direction of second-order photocurrent is strongly correlated with intrinsic ferroic orders and external light polarization. The concept of ferroicity-driven nonlinear photocurrent switching illustrated here is not limited to 2D multiferroics, rather it can be generally applicable to many ferroics (e.g. ferroelectric inorganic perovskites and hybrid organic-inorganic perovskites) and even ferroelectric metals. One can envisage to directly control nonlinear photocurrent by switching ferroelastic strain and/or ferroelectric polarization accompanied by instantaneous direction and/or sign switching of the photocurrent. Furthermore, one may conduct high-resolution characterization of ferroelastic and ferroelectric orders as well as domain evolution using ultrafast optical techniques based on ferroicity-dependent nonlinear photocurrent. The present findings thus will open up avenues for realizing configurable nonlinear optoelectronics, bulk photovoltaics, nonlinear multiferroics, etc. utilizing their ferroic orders and various nonlinear responses under external field.
6. NONLINEAR ANOMALOUS HALL EFFECT

Under broken time reversal symmetry such as in the presence of external magnetic field or internal magnetization, a transverse voltage can be established in materials perpendicular to both longitudinal current and applied magnetic field, known as classical Hall effect. However, this symmetry constraint can be relaxed in the nonlinear regime, thereby enabling nonlinear anomalous Hall current in time-reversal invariant materials an underexplored realm with exciting new opportunities beyond classical linear Hall effect. Here, using group theory and first-principles theory, we demonstrate a remarkable ferroelectric nonlinear anomalous Hall effect in time-reversal invariant few-layer WTe$_2$ where nonlinear anomalous Hall current switches in odd-layer WTe$_2$ while remaining invariant in even-layer WTe$_2$ upon ferroelectric transition. This even-odd oscillation of ferroelectric nonlinear anomalous Hall effect was found to originate from the absence and presence of Berry curvature dipole reversal and shift dipole reversal due to distinct ferroelectric transformation in even and odd-layer WTe$_2$. Our work not only treats Berry curvature dipole and shift dipole on an equal footing to account for intraband and interband contributions to nonlinear anomalous Hall effect, but also establishes Berry curvature dipole and shift dipole as new order parameters for noncentrosymmetric materials. The present findings, therefore, suggest that ferroelectric metals and Weyl semimetals may offer unprecedented opportunities for the development of nonlinear quantum electronics.

6.1 Introduction to Nonlinear Anomalous Hall Effect

In classical linear Hall effect, a transverse voltage can be developed in materials with broken time-reversal symmetry only (e.g. in the presence of external magnetic field or internal magnetization) due to Onsager’s relation. Second and higher order conductivity tensors, however, are not subject to this constraint, thereby enabling nonlinear anomalous Hall effect (NAHE) in time-reversal invariant system [27, 28, 212, 213]. NAHE was observed very recently in few-layer
tungsten ditelluride (WTe$_2$) [214–220], a layered material which also holds rich physics including high-temperature quantum spin Hall phase [74, 221–223] and electrostatic gating induced superconductivity [224, 225], in its 1T’ monolayer and type-II Weyl semimetallicity [138], large non-saturating magnetoresistance [139] and ultrafast symmetry switching [226] in its bulk phase. Monolayer 1T’ WTe$_2$ is centrosymmetric with vanishing even-order nonlinear current response, however vertical electric field can break its two-fold screw rotation symmetry, generate Berry curvature dipole (BCD), and induce second-order nonlinear anomalous Hall current [214–217]. In contrast to monolayer WTe$_2$, bilayer WTe$_2$ is naturally noncentrosymmetric due to the loss of two-fold screw rotation symmetry, resulting in intrinsic nontrivial BCD in bilayer WTe$_2$ [218–220]. Surprisingly, ferroelectric switching was recently discovered in semimetallic bilayer and few-layer WTe$_2$ [227], very unusual as ferroelectricity and semimetallicity normally do not co-exist in the same material [228]. The subtlety lies in the reduced screening along the out-of-plane direction which gives rise to finite out-of-plane ferroelectric polarization while preserving in-plane semimetallic nature. Conductance hysteresis persisting up to 300 K shows its great potential for room temperature device application. These recent studies combined reveal a striking feature of noncentrosymmetric few-layer WTe$_2$ the coexistence of ferroelectricity and NAHE within a single material, enkindling a few fundamentally and technologically important questions: what’s the fundamental correspondence between NAHE and ferroelectricity in ferroelectric metals and Weyl semimetals? Compared to ferroelectric semiconductors [229], what are the unique advantages of ferroelectric metals [227] and ferroelectric Weyl semimetals [129, 230]? Here using first-principles approach and group theoretical analysis we show an intriguing ferroelectric nonlinear anomalous Hall effect (FNAHE) in time-reversal invariant few-layer WTe$_2$. In particular, while both bilayer and trilayer WTe$_2$ possess switchable out-of-plane electronic polarization, nonlinear transverse Hall current only switches in trilayer WTe$_2$ upon ferroelectric switching. The microscopic origin of FNAHE in trilayer WTe$_2$ is found to be rooted in the reversal of Berry curvature dipole and shift dipole upon ferroelectric transition, which reveals an exciting yet unexplored realm of ferroelectric metals and Weyl semimetals with potential applications in nonlinear electronics.
6.2 Second-Order DC Current

Consider an oscillating electric field \( \mathbf{E}(r, t) = \mathbf{E}(\omega)e^{i(\mathbf{k}\cdot \mathbf{r} - \omega t)} + \mathbf{E}(-\omega)e^{-i(\mathbf{k}\cdot \mathbf{r} - \omega t)} \) with \( \mathbf{E}(\omega) = \mathbf{E}^*(-\omega) \) (e.g. under AC electric field or upon coherent light illumination), the second-order nonlinear dc current under minimal coupling approximation was derived by Sipe et al. [30], i.e.

\[
j_{a}^{0} = \chi_{abc}(0; \omega, -\omega)E_{b}(\omega)E_{c}(-\omega), \quad \text{where } \chi_{abc} \text{ are the dc photocurrent susceptibility.} \]

In general \( j_{a}^{0} \) consists of two parts depending on the polarization of electric field/incident light, including linear photogalvanic effect (LPGE) and circular photogalvanic effect (CPGE) [28, 30, 231], i.e. \( j_{a}^{0} = j_{a}^{L} + j_{a}^{C} \). BCD-induced nonlinear photocurrent current was generalized to the multiple-band case by Morimoto et al. [47] using Floquet theory and Rostami et al. [232] using density matrix beyond semiclassical Boltzmann theory. Nonlinear photocurrent originating from CPGE is also known as injection current [30].

Both LPGE and CPGE have intraband and interband contributions. For the sake of completeness we include all the terms as follows,

\[
j_{a}^{L} = j_{a, \text{intra}}^{L} + j_{a, \text{inter}}^{L}
\begin{align*}
j_{a, \text{intra}}^{L} &= -\frac{\epsilon_{abc}^{3}}{\hbar^{2}} \text{Re} \left( \frac{\tau}{1-i\omega\tau} \right) \epsilon_{adc}D_{bd}^{\text{intra}} \text{Re} \left( E_{b}(\omega)E_{c}(-\omega) \right) \\
j_{a, \text{inter}}^{L} &= -\frac{\epsilon_{abc}^{3}}{\hbar^{2}} \tau D_{a, bc}^{\text{inter}} \text{Re} \left( E_{b}(\omega)E_{c}(-\omega) \right)
\end{align*}
\]

and

\[
j_{a}^{C} = j_{a, \text{intra}}^{C} + j_{a, \text{inter}}^{C}
\begin{align*}
j_{a, \text{intra}}^{C} &= -\frac{\epsilon_{abc}^{3}}{\hbar^{2}} \text{Im} \left( \frac{\tau}{1-i\omega\tau} \right) D_{ab}^{\text{intra}} \text{Im} \left( E(\omega) \times E(-\omega) \right)_{b} \\
j_{a, \text{inter}}^{C} &= -\frac{\epsilon_{abc}^{3}}{\hbar^{2}} \tau D_{a, bc}^{\text{inter}} \text{Im} \left( E(\omega) \times E(-\omega) \right)_{b}
\end{align*}
\]

Here \( \tau \) is relaxation time and \( \epsilon_{adc} \) is the Levi-Civita symbol. \( D_{ab}^{\text{intra}} \) is the well-known BCD for intraband nonlinear process [28]. \( D_{ab}^{\text{inter}} \) is BCD for interband process associated with CPGE [218]. \( D_{a, bc}^{\text{inter}} \) is shift dipole (SD), originated from the simultaneous displacement of wavepacket upon excitation. More specifically, they are given by

\[
D_{ab}^{\text{intra}}(\mu) = \int_{BZ} f_{0}(\mu) \partial_{\alpha} \Omega_{\beta} = \int_{BZ} [dk] \sum_{n} f_{n}(\mu) \omega_{n}(k) \Omega_{\alpha\beta}^{n}(k) \delta \left( \hbar\omega_{n}(k) - \mu \right)
\]

\[
D_{ab}^{\text{inter}}(\mu, \omega) = \int_{BZ} [dk] \sum_{mn} f_{nm}(\mu) \Delta_{mn}^{a}(k) \Omega_{\beta}^{n}(k) \text{Re} \left( \frac{\tau}{1-i(\omega-\omega_{mn})\tau} \right)
\]

\[
D_{a, bc}^{\text{inter}}(\mu, \omega) = \int_{BZ} [dk] \sum_{mn} f_{nm}(\mu) R_{mn}^{a}(k) \{ r_{nm}^{b}, r_{nm}^{c} \} \text{Re} \left( \frac{1}{1-i(\omega-\omega_{mn})\tau} \right)
\]

115
Here, $\hbar \omega_n(k), v^b_n(k),$ and $f_n(\mu)$ are band energy, group velocity, and chemical-potential $\mu$ dependent Fermi-Dirac distribution, respectively. $f_{nm}(\mu) \equiv f_n(\mu) - f_m(\mu),$ and $[dk] \equiv d^d k/(2\pi)^d$ for $d$-dimension integral. $\Delta_{nm}^a \equiv v^a_n - v^a_m$ is the group velocity difference between two bands. $r^a_{nm}$ is interband Berry connection or dipole matrix element. $\Omega_{nm}^c(k)$ is the interband Berry curvature between two bands, defined as $\Omega_{nm}^c(k) \equiv i \epsilon_{abc} r^a_{nm}, r^b_{mn}.$ $\Omega_{n}^c(k)$ is the intraband Berry curvature for band $n$, given by $\Omega_{n}^c(k) = \sum_{nm} \Omega_{mn}^c(k).$ In addition, $\{ r^b_{nm}, r^c_{mn} \} \equiv r^b_{nm} r^c_{mn} + r^c_{mn} r^b_{nm},$ $R_{mn}^a$ is shift vector, given by $R_{mn}^a \equiv - \partial \phi_{mn}(k)/\partial k^a + r^a_{nm}(k) - r^a_{nn}(k),$ where $\phi_{mn}(k)$ is the phase factor of the interband Berry connection and $r^a_{nn}$ is intraband Berry connection. $\Omega_{nm}^b(k), \Omega_{n}^b(k)$ and $R_{mn}^a$ are all gauge invariant. For linearly polarized incident light/electric field, $E_b = E_c,$ hence we denote $D^{L,\text{inter}}_{ab} \equiv D_a^{\text{inter}} \text{ as well as SD } (D^{L,\text{inter}}_{ab}) \text{ have the same units of } L^{3-d} \text{ for } d\text{-dimensional system. Thus, BCD and SD have units of length in 2D, but become dimensionless in 3D.}$

The appearance of relaxation time in the dc current from interband LPGE ($j_a^{L,\text{inter}}$) seems different from the widely-used $\tau$-independent shift current formula by Sipe et al. [30], however the latter was derived for the clean limit when relaxation time $\tau$ approaches to infinite. In fact, as $\tau \rightarrow \infty, \tau \operatorname{Re} \left( \frac{1}{1-i(\omega-\omega_{mn})/\tau} \right) \rightarrow \pi \delta (\omega_{mn} - \omega),$ and the original $\tau$-time independent shift current susceptibility can be exactly recovered from the above $j_a^{L,\text{inter}}$ formula. In reality, quasiparticles do have finite relaxation time, thus $j_a^{L,\text{inter}}$ shall depend on relaxation time. The derivation of SD and interband BD is following. SC is written as $j_a^{L,\text{inter}} = 2\sigma_{abc}(0; \omega, -\omega) \operatorname{Re} (E_b(\omega) E_c(-\omega)),$ and the corresponding shift current susceptibility is given by

$$
\sigma_{2}^{abc}(0; \omega, -\omega) = -\frac{\pi e^3}{\hbar^2} \int [dk] \sum_{nm} f_{nm} R_{nm}^a(k) \{ r^b_{nm}, r^c_{mn} \} \delta (\omega_{mn} - \omega). \tag{6.4}
$$

According to the Sokhotski-Plemelj formula with a small smearing width $\epsilon = 1/\tau$ (e.g. due to finite relaxation time),

$$
\lim_{\epsilon \rightarrow 0} \frac{1}{\omega_{mn} - \omega - i\epsilon} = \mathcal{P} \frac{1}{\omega_{mn} - \omega} + i \pi \delta (\omega_{mn} - \omega), \tag{6.5}
$$

116
and $\mathcal{P}$ stands for the principal value. It leads to the following expression for the delta function,

$$
\delta(\omega_{mn} - \omega) = \frac{1}{\pi} \text{Im} \frac{1}{\omega_{mn} - \omega - i\epsilon} = \frac{1}{\pi} \text{Im} \frac{\tau}{\tau (\omega_{mn} - \omega) - i},
$$

(6.6)

By replacing the delta function with the above expression in the shift current susceptibility, we can obtain the shift current and shift dipole as follows

$$
\mathcal{J}^{L}_{a, \text{inter}} = \frac{-2e^3}{\hbar^2} \tau \mathcal{D}^{L, \text{inter}}_{a, bc} \text{Re}(E_b(\omega)E_c(-\omega)) ,
$$

(6.7)

and

$$
\mathcal{D}^{L, \text{inter}}_{a, bc}(\mu, \omega) = \int_{\text{BZ}} [dk] \sum_{nm} f_{nm}(\mu) R^a_{mn}(k) \{ r^b_{nm}, r^c_{nm} \} \text{Re} \left( \frac{1}{1 - i(\omega - \omega_{mn})\tau} \right).
$$

(6.8)

It is worth to note that in the clean limit, $\tau \to \infty$, thus $\tau \text{Re} \left( \frac{1}{1 - i(\omega - \omega_{mn})\tau} \right) \to \pi \delta(\omega_{mn} - \omega)$. The original relaxation-time independent $\sigma^{abc}(0; \omega, -\omega)$ can be exactly recovered. In reality, quasiparticles have finite relaxation time, thus $\mathcal{J}^{L}_{a, \text{inter}}$ shall depend on relaxation time. Similarly, the circular photogalvanic effect (CPGE) susceptibility is given by

$$
\eta^{a, b}_2(0; \omega, -\omega) = \frac{i\pi e^3}{2\hbar^2} \int [dk] \sum_{nm\sigma} f_{nm} \Delta^a_{mn} \Omega^b_{mn}(k) \delta(\omega_{mn} - \omega),
$$

(6.9)

which is also known as injection current susceptibility [30]

$$
\frac{dj^c_{a, \text{inter}}}{dt} = -\text{Im} \eta^{a, b}_2(0; \omega, -\omega) \text{Im}(E(\omega) \times E(-\omega))_b.
$$

(6.10)

We have

$$
\mathcal{J}^{C}_{a, \text{inter}} = \frac{e^3}{2\hbar^2} \tau \mathcal{D}^{C, \text{inter}}_{ab} \text{Im}(E(\omega) \times E(-\omega))_b.
$$

(6.11)
The Berry curvature dipole due to the interband process, given by

\[ D_{ab}^{\text{C, inter}}(\mu, \omega) = \int_{BZ} [d\mathbf{k}] \sum_{mn} f_{nm}(\mu) \Delta_{mn}^a(\mathbf{k}) \Omega_{nm}^b(\mathbf{k}) \text{Re} \left( \frac{\tau}{1 - i(\omega - \omega_{mn})} \right). \]  

(6.12)

The Berry curvature dipole due to the intraband process, is written as

\[ D_{ab}^{\text{intra}}(\mu) = \int_{BZ} f_0(\mu) \partial_a \Omega^b = \int_{BZ} [d\mathbf{k}] \sum_n f_n(\mu) v_n^a(\mathbf{k}) \Omega_n^b(\mathbf{k}) \delta(h\omega_n(\mathbf{k}) - \mu). \]  

(6.13)

Thereafter, we finally have two types of dipoles contributing to the nonlinear photocurrent shift dipole \( (D_{a,bc}^{\text{L, inter}}(\mu, \omega)) \), and Berry curvature dipole \( (D_{ab}^{\text{intra}}(\mu), D_{ab}^{\text{C, inter}}(\mu, \omega)) \). Moreover, it is worth to classify the contributions of LPGE/CPGE-induced dc current at the low/high frequency region. At the low frequency limit, \( \omega \tau \to 0 \), hence \( \tau/(1 - i\omega\tau) \to \tau \). In this case, the photocurrent due to both intraband and interband CPGE as well as interband LPGE will vanish, however a dc current from intraband LPGE will remain finite which is perpendicular to the applied electric field, thereby inducing static NAHE. At high frequency, CPGE (i.e. injection current) and interband LPGE (i.e. shift current) will have nontrivial contribution to total nonlinear photocurrent, referred as to dynamic NAHE.

It is important to note that the direction of nonlinear photocurrent induced by CPGE and LPGE have very different symmetry properties. The LPGE-induced dc current cannot flow normal to a mirror plane, however it is allowed for the CPGE-induced dc current. This distinct symmetric property of CPGE- and LPGE-induced dc current can be used to help distinguish different contributions. In fact, this is what we will see in bilayer and trilayer WTe\(_2\). More importantly, nonlinear dc current may switch their direction upon certain ferroelectric transition, giving rise to FNAHE which is the focus of this work. Below we will first reveal the fundamental difference between ferroelectric transitions in bilayer and trilayer WTe\(_2\), then demonstrate a striking even-odd anomaly of NAHE, i.e. FNAHE, in bilayer and trilayer WTe\(_2\) and provide an explanation using group theoretical analysis as well as its implication for potential FNAHE-based quantum devices.
6.3 First-Principles Tight-Binding Calculations

Here we are interested in bilayer and trilayer WTe$_2$, and both were found to exhibit ferroelectric switching. First-principles calculations for structural relaxation and electric polarization were performed using density-functional theory [87] as implemented in the Vienna Ab initio Simulation Package (VASP) [147] with the projector-augmented wave method [148]. We employed the generalized-gradient approximation of exchange-correlation functional in the Perdew-Burke-Ernzerhof form [187], a plane-wave basis with an energy cutoff of 300 eV, a Monkhorst-Pack k-point sampling of 6×12×1 for the Brillouin zone integration, and optB88-vdW functional [151] to account for dispersion interactions. Ground state crystal structures were obtained by fully relaxing both atomic positions and in-plane lattice parameters while keeping a large vacuum region of ~20 Å along the out-of-plane direction to reduce the periodic image interactions. The convergence criteria for maximal residual force was less than 0.005 eV/Å, and the convergence criteria for electronic relaxation is $10^6$ eV. We have tested higher energy cutoff of 400 eV, and the difference in the lattice constants is less than 0.04%. In addition, total electronic polarization was calculated by directly integrating the product of charge density/ionic charge and their corresponding position without using Berry phase approach. To compute the NHLE-related quantities, we first construct quasimolecular spinor Wannier functions and tight-binding Hamiltonian from Kohn-Sham wavefunctions and eigenvalues under the maximal similarity measure with respect to pseudoatomic orbitals [153, 154]. Spin-orbit coupling is taken into account, and hybrid exchange-correlation energy functional HSE06 [233] is employed with the range-separation parameter $\lambda=0.2$. In the HSE functional, the Coulomb kernel can be decomposed by complementary error function erfc($x$) and error function erf($x$), that is,

$$
\frac{1}{r} = \left. \frac{\text{erfc}(\lambda r)}{r} \right|_{\text{short-range}} + \left. \frac{\text{erf}(\lambda r)}{r} \right|_{\text{long-range}},
$$

where $\lambda$ is the range-separation parameter. $\text{erfc}(0)=\text{erf}(\infty)=1$, $\text{erfc}(\infty)=\text{erf}(0)=0$. When $\lambda$ approaches zero, the short-range part becomes the bare Coulomb operator $1/r$, hence it recovers the
pure PBE0 functional. When $\lambda$ approaches $\infty$, the long-range part dominates the Coulomb operator, the HSE functional recovers the pure PBE functional. For HSE06, $\lambda=0.2$, which is used in the calculations presented in the thesis. Total 112 and 168 quasiatomic spinor Wannier functions were obtained for bilayer and trilayer WTe$_2$, respectively. Using the developed tight-binding Hamiltonian we then compute CPGE and LPGE susceptibility tensor with a modified WANNIER90 code [234] using a dense k-point sampling of $600\times600\times1$ for both bilayer and trilayer WTe$_2$. A small imaginary smearing factor $\eta$ of 0.05 eV is applied to fundamental frequency, and Sokhotski-Plemelj theorem is employed for the Dirac delta function integration. In addition, we tested the range separation parameter $\lambda$ in the hybrid HSE functional. Although the values of Berry curvature etc. can change with respect to $\lambda$, the presence (absence) of Berry curvature switching in trilayer (bilayer) remains the same (see Figure 6.1 for HSE functional with $\lambda=0.4$). We also checked the convergence of the k-point sampling by increasing it to $1000\times1000\times1$ (Figure 6.2). Finally, since few-layer WTe$_2$ is either semimetallic or having very small gap, the dielectric screening is large, thus effect of the Coulombic interaction between electrons and holes is negligible.

![Figure 6.1: Band structure and intraband Berry curvature of bilayer WTe$_2$ using quasiatomic Wannier function with HSE functional ($\lambda=0.4$) and spin-orbit coupling. a band structure using tight-binding approach. b band structure using Kubo formula. It shows that, except some numerical difference, the Berry curvature patterns calculated using the above two methods are very similar.](image-url)
Figure 6.2: Berry curvature dipole of $-iFE$ trilayer WTe$_2$ with k-point sampling of 600×600×1 and of 1000×1000×1. It clearly shows that 600×600×1 is enough to reach the convergence, which was subsequently used for other Berry curvature, BCD and SD calculations presented in the thesis.

6.4 Ferroelectric Transition in Bilayer and Trilayer WTe$_2$

Both bilayer and trilayer WTe$_2$ were found to exhibit ferroelectric switching, however their transformation is fundamentally different, which plays a key role in their distinct NAHE. Crystal structures of monolayer, bilayer, and trilayer WTe$_2$ are shown in Figure 6.3. Monolayer 1T’ WTe$_2$ has a $C_{2h}$ point group with a mirror plane symmetry $M_y$ perpendicular to $y$-axis and a two-fold screw rotation symmetry $C_{2y}$. This leads to inversion symmetry $I = M_yC_{2y}$ or $C_{2y}M_y$. Upon van der Waals (vdW) Td stacking, multilayer noncentrosymmetric Td WTe$_2$ possess mirror plane symmetry $M_y$ only, but no longer hold $C_{2y}$ symmetry as the rotation axes of different layers are not related by any symmetry operation in the point group, consequently multilayer Td WTe$_2$ lost inversion center with $C_s$ point group.
Figure 6.3: Crystal structure of monolayer, bilayer, and trilayer WTe$_2$. (a) Monolayer 1T' WTe$_2$ with centrosymmetric C$_{2h}$ point group. It has a mirror plane $M_y$ and a screw rotation symmetry $C_{2y}$, which leads to the inversion symmetry $I = M_y C_{2y}$. (b,c) Bilayer and trilayer Td WTe$_2$ with $C_s$ point group. $C_{2y}$ symmetry is broken, hence the inversion symmetry $I$ is also broken in bilayer and trilayer WTe$_2$. 
Figure 6.4: Ferroelectric transition in bilayer and trilayer WTe$_2$. (a-c) ferroelectric transition in bilayer WTe$_2$. The two opposite ferroelectric states ($-m$FE and $+m$FE) in bilayer are transformed through a glide plane operation \{M$_{zt_a}$\}, that is, a mirror operation $M_z$ followed by an in-plane shift along $x$ by $t_a$ with $t_a = 1/2a + d_x$. The intermediate PE state $C_{2v}$ point group, thus its out-of-plane polarization with vanishes due to the glide plane \{M$_z1/2a$\} symmetry. (d-f) ferroelectric transition in trilayer WTe$_2$. The two opposite ferroelectric states ($-i$FE and $+i$FE) in trilayer are related to each other through an inversion operation \{I0\}. The intermediate PE state of trilayer WTe$_2$ has $C_{2h}$ point group, thus its out-of-plane polarization also vanishes due to inversion symmetry. The red and green vertical dashed lines show the relative shift $\pm d_x$ between adjacent WTe$_2$ layers. The corresponding in-plane shift is very small ($d_x \sim 20$ pm), therefore it is exaggerated in the above plots for illustrative purpose.
Ferroelectric transition pathways of bilayer and trilayer Td WTe$_2$ are shown in Figure 6.4. In both cases, two opposite ferroelectric (FE) states (Figure 6.4(a,c) for bilayer, Figure 6.4(d,f) for trilayer) can switch to each other by a small in-plane shift between adjacent layers along $x$ by $2d_x$ ($d_x \sim 20$ pm), passing through an intermediate paraelectric (PE) state. The intermediate PE state in bilayer WTe$_2$ (Figure 6.4(b)) has a C$_{2v}$ point group with additional $\{M_z \frac{1}{2} a\}$ symmetry, thus its out-of-plane electric polarization $P_z$ vanishes. While the ferroelectric transition is achieved by in-plane $2d_x$ shift, of two FE states are related by a glide plane operation $\{M_z t_a\}$ consisting of a mirror symmetry operation followed by a translation along $x$ by a fractional translation $t_a$ where $t_a = 1/2a + d_x$. For this reason, we denote the two FE states of bilayer WTe$_2$ by $-m$FE and $+m$FE (Figure 6.4(a,c)). In contrast, the two opposite FE states in trilayer WTe$_2$ are related by an inversion operation $I$, thus denoted by $-i$FE and $+i$FE (Figure 6.4(d,f)). Further, its intermediate PE state has a C$_{2h}$ point group with inversion symmetry, hence the out-of-plane polarization $P_z$ of the PE state in trilayer WTe$_2$ vanishes as well.

Next, we calculate total electronic polarization by summing the ionic and electronic contributions. Since we are interested in the polarization along the out of plane direction $P_z$, we can directly integrate the product between charge density/ionic charge and their corresponding position to obtain $P_z$ without using Berry phase approach. More specifically,

$$P_z = \frac{1}{S} \left( \sum_I Q_I \cdot (z_I - R_0^z) - e \int_V \rho(r) \left( z - R_0^z \right) d^3r \right), \quad (6.15)$$

where $S$ is the in-plane area of the unit cell, $Q$ is ionic charge, $\rho$ is electronic charge density, and $R_0$ is a reference point which is set to the origin of the unit cell in the present case. The equilibrium electronic charge density ($\rho$) was obtained from first-principles density functional theory (DFT) as implemented in the Vienna Ab initio Simulation Package (VASP). The calculated total electronic polarization $P_z$ is $\pm 1.67 \times 10^{-2}$ nm·µC/cm$^2$ for $\pm m$FE in bilayer, and $\pm 0.81 \times 10^{-2}$ nm·µC/cm$^2$ for $\pm i$FE in trilayer. This is in good agreement with experimentally measured vertical polarization in bilayer WTe$_2$ of $\sim 10^4$ ecm$^{-1}$ (i.e. $1.60 \times 10^{-2}$ nm·µC/cm$^2$) [227]. Additionally,
the intermediate PE state has been recently observed in experiments [226]. In brief, the results from the DFT calculations confirmed the ferroelectricity in both bilayer and trilayer WTe₂, however the symmetry relations between the two FE states are very different in the bilayer and trilayer cases, i.e. \(-m\text{FE} \leftrightarrow \text{PE} \leftrightarrow +m\text{FE}\) and \(-i\text{FE} \leftrightarrow \text{PE} \leftrightarrow +i\text{FE}\), which is essential for understanding their distinct NAHE upon ferroelectric switching we will discuss shortly. In addition, as the polarization depends on the interlayer distance \((d_z)\), we conducted a series of polarization calculations by changing the interlayer distance with respect to the equilibrium configuration by \(\Delta d_z\). The calculated result is shown below. In addition, the corresponding total energy is plotted as a function of vertical polarization, as shown in Figure 6.5. It clearly shows that the total electric polarization has significant increase when interlayer distance reduces.

![Figure 6.5: Electric polarization of bilayer WTe₂. a Out-of-plane polarization of bilayer WTe₂ as a function of distance away from equilibrium configuration, \(\Delta d_z\), between two layers. b Total energy as a function of the out-of-plane polarization.](image)

6.5 NAHE in Bilayer and Trilayer WTe₂ upon Ferroelectric Switching

Now we proceed to discuss NAHE in few-layer WTe₂, in particular ferroelectric switching of NAHE (i.e. FNAHE) in odd-layer WTe₂, and reveal the intriguing connection between BCD/SD and ferroelectric order. We compute their electronic structure by first-principles DFT using hybrid exchange-correlation functional with spin-orbit coupling taken into account. Quasitatomic spinor
Wannier functions and tight-binding Hamiltonian were obtained by rotating and optimizing the Bloch functions with a maximal similarity measure with respect to pseudoatomic orbitals [153, 154]. Subsequently, first-principles tight-binding approach was applied to compute all the physical quantities such as band structures, BCD, SD, Berry curvature etc.

Electronic band structure of bilayer WTe$_2$ is presented in Figure 6.6(a), color-coded by total intraband Berry curvature of all occupied bands, that is, \( \Omega^{\text{z,intra}}_{\alpha\beta}(k) = \sum_n f_n \Omega^{\text{z}}_n(k) \). It shows bilayer WTe$_2$ is a small gap insulator, and the intraband Berry curvature is odd with respect to due to the presence of time-reversal symmetry. The k-dependent intraband Berry curvature \( \Omega^{\text{z,intra}}_{\alpha\beta}(k) \) are shown in Figure 6.6(c,d) at two different chemical potentials of \( \mu = \pm 50 \text{ meV} \). Alternatively, one may use the Kubo formula with the sum-over-states approach for Berry curvature. Intraband Berry curvature reads

\[
\Omega^{\gamma}_{\alpha\beta}(k) = \epsilon_{\alpha\beta\gamma} \Omega^{\gamma}_{\alpha\beta}(k),
\]

where

\[
\Omega^{\gamma}_{\alpha\beta}(k) = - \text{Im} \left( \frac{\sum_{m\neq n} v^{\alpha}_{nm}(k) v^{\beta}_{mn}(k)}{\omega^2_m(k) - \omega^2_n(k)} \right).
\]

Here we consider the \( z \) component

\[
\Omega^{z}_{\alpha\beta}(k) = - \text{Im} \left( \frac{\sum_{m\neq n} v^{x}_{nm}(k) v^{y}_{mn}(k) - v^{y}_{nm}(k) v^{x}_{mn}(k)}{\omega_m(k) - \omega_n(k)} \right).
\]

Similarly, interband Berry curvature \( \Omega^{z}_{\alpha\beta}(k) \) with at frequency \( \omega = 120 \text{ meV} \) is displayed in Figure 6.6(e,f) for two sets of occupied and unoccupied bands around the Fermi energy, \( \Omega^{z,\text{inter}}_{\text{VBM-1,CBM}}(k) \) and \( \Omega^{z,\text{inter}}_{\text{VBM,CBM-1}}(k) \), respectively. VBM refers to valence band maximum, and CBM refers to conduction band minimum. The Berry curvature distribution plots confirm the presence of mirror symmetry \( M_y \) and time-reversal symmetry \( T \). Thus, the integral of the intraband Berry curvature over the full Brillouin zone vanishes, and linear anomalous Hall effect is absent.
Figure 6.6: Electronic structure and NAHE of bilayer WTe$_2$. (a) Band structure of bilayer WTe$_2$ in ±mFE state color-coded by the $z$-component of intraband Berry curvature $\Omega_z^{\text{intra}}(k)$. Spin-orbit coupling is included and hybrid HSE06 functional is employed. (b) BCD and SD tensor elements $D_{yz}^{\text{intra}}(\mu)$, $D_{yz}^{\text{C,inter}}(\mu, \tilde{\omega})$, and $D_{xy}^{L,\text{inter}}(\mu, \tilde{\omega})$ as function of chemical potential $\mu$. For interband BCD and SD, $\tilde{\omega}$ is set to 120 meV. (c,d) $k$-dependent distribution of intraband Berry curvature $\Omega_z^{\text{intra}}(k)$ at $\mu = \pm50$ meV, respectively. (e,f) $k$-dependent distribution interband Berry curvature $\Omega_z^{\text{inter}}(k)$ for $(n,m)=$(VBM-1,CBM) and for $(n,m)=$(VBM-1,CBM), respectively. The results are the same for both +mFE and -mFE state, suggesting that nonlinear anomalous Hall current can be induced in bilayer WTe$_2$, but it will not switch sign upon ferroelectric transition.

Furthermore, Figure 6.6(b) shows the calculated BCD and SD tensor elements $D_{yz}^{\text{intra}}$, $D_{yz}^{\text{C,inter}}$, and $D_{xy}^{L,\text{inter}}$ the key physical quantities governing NAHE. It clearly demonstrates the presence of finite BCD and SD and thus NAHE in bilayer WTe$_2$. The calculated BCD varies between 0-0.4 Å depending on the chemical potential, which is in nice agreement with the experimental values of...
0.1-0.7Å by Kang et al. [219]. Moreover, upon ferroelectric transition between -$m$FE and +$m$FE state, the Berry curvature, BCD and SD remain unchanged, thus nonlinear anomalous Hall current will not switch direction upon ferroelectric transition in bilayer WTe$_2$. Similarly, the out-of-plane spin polarization remains unflipped, while the in-plane spin polarization is expected to reverse. Furthermore, Du et al. recently studied NAHE in bilayer WTe$_2$ using a model Hamiltonian and found that, as the SOC strength evolves, BCD becomes strong near titled band anticrossings and band inversions [220]. Our first-principles results also show large Berry curvature near band anticrossing which is consistent with the conclusion from Du et al.’s analysis. The magnitude of the calculated Berry curvature is similar to that in Ma et al. [218] for bilayer WTe$_2$ in the absence of electric field. The difference in the detailed band structure is mainly due to the electronic structure sensitive to DFT exchange-correlation functional, vdW functional, and the Wannier function construction. Nevertheless, both our results and the work by Ma et al. [218] show the nontrivial BCD contribution to NAHE.

Trilayer WTe$_2$ is quite different from bilayer WTe$_2$. Figure 6.7(a,b) show its electronic band structure of -$i$FE and +$i$FE state, respectively. In contrast to the bilayer case, intraband Berry curvature changes sign upon ferroelectric transition. The similar sign change is also evidenced in the opposite $k$-dependent intraband and interband Berry curvature $\Omega_{\text{occ}}^{z,\text{intra}}(k)$ and $\Omega_{\text{VBM-1,CBM}}^{z,\text{inter}}(k)$ as displayed in Figure 6.7(e-h). Consequently, the sign of BCD and SD flips upon ferroelectric transition between -$i$FE and +$i$FE, demonstrated in Figure 6.7(c,d). Therefore, in direct contrast to bilayer WTe$_2$, the nonlinear dc current in trilayer WTe$_2$ will switch its direction upon ferroelectric transition. The calculated BCD ranges from 0 to 0.7Å depending on the chemical potential, also in good agreement with experiment [219]. Moreover, there is a clear plateau in $D_{y2}^{C,\text{inter}}$ marked by purple arrow in Figure 6.7(c). It is originated from the large joint density of state around 120 meV indicated by purple arrow in Figure 6.7(a), which remains constant when the chemical potential is located between the energy window. It is also worth to note that, like the bilayer case, the integral of Berry curvature of trilayer WTe$_2$ is also zero due to the presence of time-reversal symmetry, hence the linear anomalous Hall effect is absent.
Figure 6.7: Electronic structure and FNAHE of trilayer WTe$_2$. (a,b) Band structure of trilayer WTe$_2$ in -$i$FE and +$i$FE state, respectively. Both are color-coded by the $z$-component of intraband Berry curvature $\Omega^z$$_{\text{intra}}$(k). (c,d) BCD tensor elements $D^\text{intra}_{yz}(\mu)$ and $D^\text{C,inter}_{yz}(\mu, \omega)$, and SD tensor element $D^\text{L,inter}_{xy}(\mu, \omega)$ as function of chemical potential $\mu$ for -$i$FE and +$i$FE state, respectively. For interband BCD and SD, $\omega$ is set to 120 meV. (e,f) k-dependent distribution of intraband Berry curvature $\Omega^z$$_{\text{intra}}$(k) at $\mu = \pm 50$ meV for -$i$FE and +$i$FE state, respectively. (g,h) k-dependent distribution interband Berry curvature $\Omega^z$$_{\text{inter}}$(k) between (VBM-1,CBM) around the Fermi surface for -$i$FE and +$i$FE state, respectively. The results clearly show that nonlinear anomalous Hall current in trilayer WTe$_2$ will switch sign upon ferroelectric transition, in direct contrast to the bilayer case.
Finally, the dc current susceptibility of bilayer and trilayer WTe$_2$ will be reversed in the trilayer case only. Figure 6.8 shows the interband LPGE susceptibility $\sigma_{abc}$ of bilayer and trilayer WTe$_2$ at $\mu=0$, which is about ten times higher than that in monolayer group IV monochalcogenides [229]. It is clear that in bilayer WTe$_2$ the two independent susceptibility tensor elements $\sigma_{xxx}$ and $\sigma_{xyy}$ of the $\pm m$FE states remain invariant upon ferroelectric transition, while for trilayer WTe$_2$ both $\sigma_{xxx}$ and $\sigma_{xyy}$ of the $\pm i$FE states flip the sign. The above electronic structure results demonstrate a striking difference between bilayer and trilayer WTe$_2$, that is, nonlinear anomalous Hall current flips its direction upon ferroelectric switching in trilayer WTe$_2$, but remains unchanged in bilayer WTe$_2$.

![Figure 6.8: Interband LPGE susceptibility $\sigma_{abc}$ of bilayer and trilayer WTe$_2$ at $\mu=0$. (a) $\sigma_{abc}$ for the $\pm m$FE states of bilayer WTe$_2$. (b) $\sigma_{abc}$ for the $\pm i$FE state of trilayer WTe$_2$. These two plots demonstrate that, under ferroelectric transition, the sign of LPGE-induced dc current susceptibility switches in the trilayer case, but remains invariant in the bilayer case.]()

6.6 Group Theoretical Analysis of NAHE in Bilayer and FNAHE in Trilayer WTe$_2$

Here we provide a group theoretical analysis of NAHE in addition to the above first-principles calculations. Both bilayer and trilayer WTe$_2$ have $C_s$ point group with a mirror symmetry $M_y$. For circularly polarized incident light propagating along $z$, $(E(\omega) \times E(-\omega))_z$, shares the same $A''$ representation as axial vector $R_z$. Therefore, $\Gamma_{jy} \otimes \Gamma_{R_{zz}} = A'' \otimes A'' = A'$, suggesting $\Gamma_{jy} \otimes \Gamma_{R_{zz}}$
includes total symmetric irreducible representation, and hence the nonlinear CPGE current can be induced along y, i.e. perpendicular to the \(xz\) mirror plane. Furthermore, \(\Gamma_{j_x,z} \otimes \Gamma_{R_{x,z}} = A''\), thus no CPGE current can be induced along \(x\). In contrast, for linearly polarized incident light/electric field with in-plane polarization, we have \(\Gamma_{j_x} \otimes \Gamma_{E_x} \otimes \Gamma_{E_x} = A' \otimes A' \otimes A' = A'\), and \(\Gamma_{j_y} \otimes \Gamma_{E_y} \otimes \Gamma_{E_y} = A' \otimes A'' \otimes A'' = A'\), indicating that the LPGE current can be induced along \(x\). However, \(\Gamma_{j_y} \otimes \Gamma_{E_x} \otimes \Gamma_{E_x} = \Gamma_{j_y} \otimes \Gamma_{E_y} \otimes \Gamma_{E_y} = A'\), thus no LPGE current can be induced along \(y\). This leads to a contrasting CPGE- and LPGE-based nonlinear anomalous Hall current in few-layer WTe\(_2\) with \(Cs\) point group, that is, linearly polarized light/electric field with in-plane polarization will generate nonlinear anomalous Hall current along \(x\) only \((j^L_x \neq 0, j^L_y = 0)\), while circularly polarized light propagating along \(z\) axis will generate nonlinear anomalous Hall current along \(y\) only \((j^C_x = 0, j^C_y \neq 0)\).

The correlation between the irreducible representations of parent group \(C_{2h}\) and its noncentrosymmetric subgroups \(C_2, C_s, \text{and } C_1\) is summarized in Table 6.1. \(C_{2h}\) has four representations and two independent symmetries \(M_y\) and \(C_{2y}\). Breaking one or two of them leads to three noncentrosymmetric subgroups. We start from monolayer 1T’ WTe\(_2\) which has point group of \(C_{2h}\), whose second order nonlinear current response vanishes due to the presence of inversion symmetry. Upon vdW stacking (e.g. few-layer and bulk Td WTe\(_2\)), \(C_{2y}\) is broken with \(M_y\) left unchanged, which breaks the inversion symmetry and results in subgroup \(C_s\). Consequently, as we analyzed above, \(j^C_x = 0\), but \(j^C_y \neq 0\) under circularly polarized light, while \(j^L_x \neq 0\) but \(j^L_y = 0\) under linearly polarized light/electric field with in-plane polarization. However, if \(M_y\) is broken with \(C_{2y}\) being preserved, it will fall into subgroup \(C_2\). In this case, \(j^C_x \neq 0\) and \(j^C_y = 0\) under circularly polarized light, while \(j^C_x = 0\) and \(j^C_y \neq 0\) under linearly \(x/y\)-polarized light/electric field. Further, if both \(M_y\) and \(C_{2y}\) are broken, it will end up with subgroup \(C_1\), and enable all possible LPGE and CPGE current responses along different directions. We now discuss the fundamental difference of NAHE between bilayer and trilayer WTe\(_2\) upon ferroelectric switching. A general symmetry operator in Seitz notation is given by \(g = \{Rt_R\}\), where \(R\) is point group symmetry operation, \(t_R\) is a translational vector. A time-reversal antisymmetric pseudovector (e.g. Berry curvature and spin polarization)
transforms under operator $g$ as follows, $m'(k) = gm(k) = P_R P_T Rm(k)$, where $P_R$ and $P_T$ are spatial and temporal parity associated with $g$, respectively. $P_T = \pm 1$ when $Rk = \pm k + k$, where $k$ is multiples of reciprocal lattice vector. For bilayer WTe$_2$, as aforementioned, two ferroelectric states can be related by a glide plane operation $M_z t_0$, where $t_0$ refers to a fractional translation along $x$. Thus, $P_R = -1$, $P_T = 1$, and $(m_x, m_y, m_z)^{+\text{FE}} = (-m_x, -m_y, m_z)^{-\text{FE}}$. For trilayer WTe$_2$, the two ferroelectric states are related by an inversion operation $\{I0\}$, thus $P_R = P_T = -1$, subsequently $(m_x, m_y, m_z)^{\text{FE}} = (-m_x, -m_y, -m_z)^{-\text{FE}}$. The above two conclusions are applicable to any time-reversal antisymmetric pseudovectors such as Berry curvature and spin polarization.

For example, for intraband and interband Berry curvature, $M_z \Omega^z(k_x, k_y) = \Omega^z(k_x, k_y)$ in bilayer WTe$_2$, and $\mathcal{I} \Omega^z(k_x, k_y) = \Omega^z(-k_x, -k_y) \stackrel{\text{TRI}}{\longrightarrow} -\Omega^z(k_x, k_y)$ in trilayer WTe$_2$, indicating that the sign of intraband and interband BCD ($D_{ab}^{\text{intra}}, D_{ab}^{\text{inter}}$) flips only in trilayer WTe$_2$ upon ferroelectric transition. This is in excellent agreement with the first-principles calculations shown in Figure 6.6(c-f) and Figure 6.7(e-h). In addition, the in-plane spin polarization switches in both cases, and the out-of-plane spin polarization becomes reversed in trilayer WTe$_2$ while remaining unflipped for bilayer WTe$_2$, which also agrees with the calculations. Figure 6.9 and Figure 6.10 show the band structure of bilayer WTe$_2$ color-coded by spin polarization $S^x_n(k)$ and $S^z_n(k)$ using the PBE functional. It’s clear to see that $S^x_n(k)^{-\text{FE}} = -S^x_n(k)^{+\text{FE}}$ and $S^z_n(k)^{-\text{FE}} = S^z_n(k)^{+\text{FE}}$. However, in trilayer case as shown in Figure 6.11, we have $S^z_n(k)^{-\text{FE}} = -S^z_n(k)^{+\text{FE}}$. Again, the first-principles calculations agree well with group theoretical analysis.

<table>
<thead>
<tr>
<th>$C_{2h}$</th>
<th>$C_2$ ($M_y$ broken)</th>
<th>$C_s$ ($C_{2y}$ broken)</th>
<th>$C_1$ (both $M_y$ and $C_{2y}$ broken)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_g(R_y)$</td>
<td>$A(y, R_y)$</td>
<td>$A'(x, z, R_y)$</td>
<td>$A(\text{all})$</td>
</tr>
<tr>
<td>$B_g(R_x, R_z)$</td>
<td>$B(x, z, R_x, R_z)$</td>
<td>$A''(y, R_x, R_z)$</td>
<td>$A$</td>
</tr>
<tr>
<td>$A_u(y)$</td>
<td>$A(y, R_y)$</td>
<td>$A''(y, R_x, R_z)$</td>
<td>$A$</td>
</tr>
<tr>
<td>$B_u(x, z)$</td>
<td>$B(x, z, R_x, R_z)$</td>
<td>$A'(x, z, R_y)$</td>
<td>$A$</td>
</tr>
</tbody>
</table>

Table 6.1: The correlation between the irreducible representations of parent group $C_{2h}$ and its noncentrosymmetric subgroups $C_2$, $C_s$, and $C_1$. 
Different from pseudovectors, polar vector \( p \) such as electric polarization and shift vector transforms as follows: \( p' = Rp \). Therefore, both mirror \( \mathcal{M}_z \) and inversion \( \mathcal{I} \) operation will lead to vertical polarization reversal, i.e. \( p'_z = \mathcal{M}_zp_z = -p_z \) and \( p'_z = \mathcal{I}p_z = -p_z \), i.e. the out-of-plane electric dipole flips sign in both bilayer and trilayer WTe\(_2\) upon ferroelectric transition. In addition, for
in-plane shift vector $R_{mn}^a$ with $a \in \{x, y\}$, $(R_{mn}^a)' = \mathcal{M}_z R_{mn}^a = R_{mn}^a$, and $(R_{mn}^a)' = \mathcal{I} R_{mn}^a = -R_{mn}^a$, indicating that the in-plane shift vector $R_{mn}^a$ and thus SD $D_{y/2}^{L, \text{inter}}$ will flip only in trilayer WTe$_2$ upon ferroelectric transition. Consequently, the total $j_x^L$ and $j_y^C$ from CPGE and LPGE will switch direction upon ferroelectric transition, provoking FNAHE in time-reversal invariant semimetals. Moreover, it suggests that the BCD and SD can serve as distinct order parameters for noncentrosymmetric semimetals. Figure 6.12(a) presents an illustrative summary of the transformation of Berry curvature, spin polarization, and electric polarization under different symmetry operation, while Figure 6.12(b,c) show the ferroelectric switching of nonlinear current in the $-i$FE and $+i$FE state of trilayer WTe$_2$. Upon the out-of-plane polarization switching, nonlinear Hall current $j_x^L$ generated via LPGE switches between $-x$ and $+x$ direction under the same external electric field with in-plane linear polarization. Moreover, nonlinear Hall current $j_x^C$ induced by CPGE switches between $-y$ and $+y$ direction under circularly-polarized light with normal incidence. Its worth to emphasize that the intermediate PE state in bilayer and trilayer WTe$_2$ has noncentrosymmetric $C_{2v}$ and centrosymmetric $C_{2h}$ point group, respectively. Thus, despite that the out-of-plane electric polarization vanishes in both cases, nonlinear anomalous Hall current of the PE state vanishes in trilayer, but remains finite in bilayer.

Figure 6.11: Band structure of trilayer WTe$_2$ color-coded by spin polarization $S^z_n(k)$ using the PBE functional. a band structure for the $-i$FE state. b band structure for $+i$FE state. Spin-orbit coupling is included. $S^z_n(k)_{-iFE} = -S^z_n(k)_{+iFE}$ holds for arbitrary k point.
Figure 6.12: Transformation of pseudovectors and polar vectors under different symmetry operation and ferroelectric switching of LPGE and CPGE nonlinear current. (a) Transformation of Berry curvature, spin polarization, shift vector, BCD and SD as well as LPGE and CPGE nonlinear current under different symmetry operation between two ferroelectric states in time-reversal invariant few-layer WTe$_2$. Berry curvature and spin polarization transform as time-reversal antisymmetric pseudovectors. Under a mirror symmetry operation $M_z$ for the ferroelectric states in bilayer WTe$_2$, most quantities remain invariant, except in-plane spin and Berry curvature component. Under inversion symmetry operation I for the -$i$FE and +$i$FE state states in trilayer WTe$_2$, all quantities, including Berry curvature, spin polarization, shift vector, BCD and SD, flip the sign in the presence of time-reversal symmetry, giving rise to FNAHE in trilayer WTe$_2$. (b,c) Ferroelectric switching of nonlinear current in the -$i$FE and +$i$FE state of trilayer WTe$_2$, respectively. Upon the polarization $P_z$ switching, nonlinear anomalous Hall current $j_{Lx}^L$ from LPGE switches between -$x$ and +$x$ direction under external field with in-plane linear polarization, while nonlinear anomalous Hall current $j_{y}^C$ from CPGE switches between -$y$ and +$y$ direction under circularly-polarized light with normal incidence.
The present work considers intrinsic NAHE due to BCD. Disorder however can play an important role in NAHE as pointed out by Du et al. [235] and Isobe et al. [236] particularly in the dc limit due to side jump and skew scattering [41]. As disorder scattering depends on scattering potential and defect density, further experimental studies are required to understand the nature of defects in bilayer and trilayer WTe$_2$. The transformation behavior of BCD-induced NAHE in multilayer WTe$_2$ upon ferroelectric transition may be utilized to distinguish itself from the disorder scattering-induced NAHE. For example, upon ferroelectric transition, the change in defect scattering potential may behave very differently from the change in crystal structure, thereby potentially helping differentiate the two NAHE contributions. In conclusion, using first-principles calculations and group theoretical analyses we investigated the NAHE in bilayer and trilayer WTe$_2$ and, more importantly, the underlying microscopic origin of FNAHE (i.e. ferroelectric switching of NAHE) in trilayer WTe$_2$. Although both bilayer and trilayer WTe$_2$ exhibit ferroelectric transition with similar electric polarization, they behave very differently in NAHE. In the trilayer case, the nonlinear anomalous Hall current flips direction upon ferroelectric switching due to the reversal of BCD and SD under an effective inversion operation of the two ferroelectric states. In contrast, the two ferroelectric states in bilayer WTe$_2$ are related effectively by a glide plane operation which does not flip the BCD/SD, thus its nonlinear anomalous Hall current will not flip upon ferroelectric switching. In addition, NAHE is expected to vanish in the PE state of trilayer WTe$_2$, but remains nontrivial for the PE state of the bilayer case. The above conclusions are applicable to any even and odd layer WTe$_2$ (except monolayer 1T’ WTe$_2$ as it is centrosymmetric with vanishing second order NAHE) as long as the two opposite ferroelectric states have the same relationship as the bilayer and trilayer case. The theoretical approaches presented here can also be applied to other materials such as Weyl semimetals [129, 230]. More importantly, our results imply that BCD and SD can serve as new order parameters for noncentrosymmetric materials, which opens up the possibility to explore nonlinear multiferroicity based on the coupling of BCD/SD and ferroelectric order. Ferroelectric metals may be advantageous as their vanishing bandgap will not only bring intraband contributions to nonlinear anomalous Hall current that is absent in semiconduc-
tors/insulators, but also significantly enhance the interband contributions due to the reduced gap of nonlinear interband processes. For example, the calculated nonlinear anomalous Hall current from interband LPGE in bilayer and trilayer WTe$_2$ is about one order magnitude higher than that in ferroelectric GeS [229]. Moreover, FNAHE provides a facile approach for direct readout of ferroelectric states, which, combined with vertical ferroelectric writing, may allow for realizing nonlinear multiferroic memory. In addition, the distinct ferroelectric transformation pathway may provide potential routes to realizing non-abelian reciprocal braiding of Weyl nodes [237]. The present findings therefore present an underexplored realm beyond classical linear Hall effect and conventional ferroelectrics with exciting new opportunities for FNAHE-based nonlinear quantum electronics using ferroelectric metals and Weyl semimetals.
7. FUTURE DIRECTIONS AND CONCLUSIONS

The results presented in the previous chapters have demonstrated the rich NLO responses in various types of materials, ranging from 2D materials to topological materials. However, there are many other interesting NLO phenomena that have not been extensively explored or discovered in the past. Below we will highlight a few exciting directions and important issues to be addressed, including surface SHG, inverse Faraday effect, optical rectification, nonlinear magnetotransport responses, nonlinear current responses under external magnetic field, higher order NLO effect, and many-body effect.

7.1 Surface Second Harmonic Generation

Surface SHG can help uncover the surface and interface electronic structure with some advantages compared to the conventional surface spectroscopies such as X-ray photoelectron spectroscopy, Auger electron spectroscopy, electron energy loss spectroscopy, etc. Here we develop a first-principles tight-binding approach to compute surface SHG efficiently.

The general second-order electronic polarization is related to the applied electric field

\[ P_a (\omega_\beta + \omega_\gamma) = \chi^{abc} (-\omega_\beta - \omega_\gamma; \omega_\beta, \omega_\gamma) E_b (\omega_\beta) E_c (\omega_\gamma). \] (7.1)

The SHG susceptibility can be separated into interband term and mixed intraband term, which is given by [30, 206, 238]

\[ \chi^{abc} (-2\omega; \omega, \omega) = \chi_e^{abc} (-2\omega; \omega, \omega) + \chi_i^{abc} (-2\omega; \omega, \omega), \] (7.2)

where

\[ \chi_e^{abc} (-2\omega; \omega, \omega) = \frac{e^3}{\hbar^2} \sum_{nmli} \int \frac{dk}{(2\pi)^d} \frac{r_{nm}^a \{ r_{ml}^b, r_{ln}^c \}}{\omega_{ml} - \omega_{ln}} \times \left( \frac{2f_{nm}}{\omega_{nm} - 2\omega} + \frac{f_{ml}}{\omega_{ml} - \omega} + \frac{f_{ln}}{\omega_{ln} - \omega} \right), \] (7.3)

138
\[
\chi_i^{abc}(-2\omega; \omega, \omega) = -\frac{ie^3}{\hbar^2} \sum_{nm} \int \frac{d\mathbf{k}}{(2\pi)^d} f_{nm}\left(\frac{1}{\omega_{mn} - \omega} - \frac{4}{\omega_{mn} - 2\omega}\right) r_{mn}^a \left\{ \frac{\Delta_{nm}^b r_{mn}^c}{\omega_{mn}^2} \right\} \\
+ \left[ \frac{\{r_{nm;\alpha}^b, r_{mn}^c\}}{\omega_{mn} (\omega_{mn} - \omega)} - \frac{r_{mn}^a (r_{mn;\alpha}^b + r_{mn;\beta}^c)}{\omega_{mn} (\omega_{mn} - 2\omega)} - \frac{(r_{nm;\alpha}^b r_{mn}^c + r_{nm;\beta}^b r_{mn}^c)}{2\omega_{mn} (\omega_{mn} - \omega)} \right].
\]

(7.4)

The mixed intraband term \(\chi_i^{abc}\) includes the contribution from the modulation of the linear susceptibility by the intraband motion and the modification of the intraband motion by the polarization energy related to the interband motion [189]. By substituting the modified position operator into the SHG calculation one can avoid destructive interference of the SHG polarizations possibly induced by the two surfaces of the finite slab, and thus the SHG response from a particular surface can be readily obtained [239]. We assume that electric field of incident light has a \(z\) dependence, then the vector potential in length gauge is given by

\[
V_A = e \left[ S(z) \mathbf{E}^{sf} \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{E}^{sf} S(z) \right] / 2.
\]

(7.5)

The emission of second harmonic light is described by the modified position operator

\[
\tilde{\mathbf{r}} = [S(z) \mathbf{r} + \mathbf{r} S(z)],
\]

(7.6)

where \(S(z) = \cos^2\left(\frac{\pi z}{2L}\right)\) is a function of \(z\), being 1 at the top surface and 0 at the bottom surface, which avoids the spurious destructive interference of second harmonic light generated at the two surfaces of the slab. Here we used Wannier functions based tight binding Hamiltonian to calculate the surface SHG susceptibilities. The modified dipole matrix elements are given by

\[
\langle n\mathbf{R}|\tilde{\mathbf{r}}|n'\mathbf{R}' \rangle = \{\{S(z_n \mathbf{R}) + S(z_{n'} \mathbf{R}')\} / 2\} \times \langle n\mathbf{R} | \mathbf{r} | n' \mathbf{R}' \rangle,
\]

(7.7)
where

\[ z_{n\mathbf{R}} = (n\mathbf{R}|z|n\mathbf{R}) \]  

(7.8)

where the matrix elements of \( S(z) \) between different atoms are ignored. For centrosymmetric materials like anatase TiO\(_2\), bulk SHG vanishes but surface SHG is nonvanishing due to the breaking of inversion symmetry due to the discontinuity, reconstruction, or relaxation of the surface [240].

![Figure 7.1: Slab model of 1L, 2L, 7L anatase TiO\(_2\) (101) surface.](image)

Here we focus on thin dielectric materials. The surface structure of 1L, 2L, 7L anatase TiO\(_2\) is shown in Figure 7.1. We calculated the surface SHG in layered anatase TiO\(_2\) and bulk SHG in GaAs as shown in Figure 7.2. The calculated \( \chi^{xxx} \) component of surface SHG is quite large, and the effective strength at the zero frequency \( \chi^{xxx} \) limit is 290 pm/V, which is comparable to GaAs. To understand the underlying mechanism, we notice that the interband dipole product \( r_{nm}^a r_{ml}^b r_{ln}^c \) plays a dominating role in the surface SHG susceptibility tensor \( \chi^{abc}(-2\omega;\omega,\omega) \). We therefore analyzed the interband dipole strength in the first Brillouin zone, and it turns out that the dipole strength between valence band maximum (VBM) and VBM-1, \( |r_{(VBM,VBM-1)}^{VBM}(k)|^2 \), is particularly large in the first Brillouin zone. As VBM and VBM-1 are close to the band edge, their large dipole strength will have a significant contribution to the surface SHG response. It is worth to mention
Figure 7.2: Imaginary and real part SHG in 7L TiO$_2$ and bulk GaAs. The effective thickness of layered TiO$_2$ included $\sim 15$ Å of vacuum. The SHG susceptibility component in GaAs is $\chi^{xyz}_2$.

Figure 7.3: Interband dipole transition strength in the first Brillouin zone. VBM denotes valence band maximum and CBM represents conduction band minimum.

that, as both VBM and VBM-1 belong to valence manifold, $|r^{\tau}_{(VBM,VBM-1)}|^2$ does not contribute to linear optical absorption. The above results thus reveal that the microscopic origin of large surface SHG comes from two factors: a) large intrinsic interband dipole matrix element of TiO$_2$ and b) surface-induced inversion symmetry breaking. The results open an avenue for exploring centrosymmetric materials as surface SHG applications.
7.2 Inverse Faraday Effect and Optical Rectification

The magneto-optical Faraday effect (MOFE) refers to the rotation of transmitted light polarization when it traverse through a magnetic medium. Kerr observed the similar rotation of reflected light from a magnetized surface, that is, magneto-optic Kerr effect (MOKE). The complex polar Kerr effect is defined by [241]

\[ \Phi = \phi_{Kerr} + i \theta_{Kerr}, \]  

(7.9)

where \( \phi \) is the rotation of the plane of polarization of light traversing the material and \( \theta \) is arising from the causality. For \( M \parallel z \), the linear optical response is described by a dielectric tensor

\[ \varepsilon = \begin{bmatrix} \epsilon_0 + \epsilon_2 & +i\epsilon_1 & 0 \\ -i\epsilon_1 & \epsilon_0 - \epsilon_2 & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix}. \]  

(7.10)

To the first order, \( \epsilon_1 \) is linear in magnetization \( M \) (classical) and the rest are independent of \( M \). The normal modes are related to the eigenvalues

\[ n_\pm^2 = \epsilon_0 \pm (\epsilon_1^2 + \epsilon_2^2)^{1/2}. \]  

(7.11)

In nonpleochroic media with \( \epsilon_2 = 0 \), e.g. three-fold rotation symmetry along \( z \), the complex polar Kerr rotation angle is given by [242]

\[ \Phi_{Kerr} = -i \epsilon_1/\epsilon_0^{1/2} (\epsilon_0 - 1). \]  

(7.12)

Similarly, the complex Faraday rotation angle is

\[ \Phi_{Faraday} = -\pi (L/\lambda) \left( \epsilon_1/\epsilon_0^{1/2} \right), \]  

(7.13)
where $L$ is the optical path length. Comparing Kerr and Faraday effects, one concludes that the Kerr effect may be derived from Faraday effect with an effective length

$$L_{\text{eff}} \approx \left(\frac{\lambda}{\pi}\right) \left|\varepsilon_0 - 1\right|^{-1}. \tag{7.14}$$

Hence Faraday effect is generally stronger than Kerr effect. For 2D materials, the optical transmission measurements are quite straightforward, and in this case Kerr effect will be comparable with Faraday effect due to the reduction of dimensionality. Previously, we studied the nonlinear electric polarization induced by light. However, the polarized light may also induce a magnetic polarization in the media. Such phenomenon was first observed in the 1960s and known as inverse Faraday effect (IFE).

The second-order correction to the magnetization is

$$\langle M^{(2)} \rangle = \text{Tr}(M\rho^{(2)}), \tag{7.15}$$

where $M = \mu_B (g_s S + L)$ is the operator for the total spin and orbit moment. The induced static magnetization is given by [243]

$$M^{\text{IFE}} = \left( M^{\text{IFE}}_o + M^{\text{IFE}}_{dA} + M^{\text{IFE}}_{dB} + \text{c.c.} \right) E_0^2, \tag{7.16}$$

where

$$M^{\text{IFE}}_o = e^2 \sum_{n \neq m; l} M_{mn} \frac{r^+_{nl} r^-_{ln} (\rho_{nm}^0 - \rho_{ln}^0)}{E_l - E_n + i\hbar \Gamma_{ln} - \hbar \omega} - \left(\frac{r^+_{nl} r^-_{ln} (\rho_{nl}^0 - \rho_{nm}^0)}{E_n - E_l + i\hbar \Gamma_{nl} - \hbar \omega}\right), \tag{7.17}$$

$$M^{\text{IFE}}_{dA} = e^2 \sum_{nl} M_{nn} \left( \frac{r^+_{nl} r^-_{ln} (\rho_{nl}^0 - \rho_{nm}^0)}{(E_l - E_n + i\hbar \Gamma_{ln} - \hbar \omega)^2} \right), \tag{7.18}$$

$$M^{\text{IFE}}_{dB} = e^2 \sum_{nl} \frac{M_{nn} r^+_{nl} r^+_{ln} (\rho_{mn}^0 - \rho_{nl}^0) (i\hbar \Gamma_{ln} - \hbar \omega)}{\hbar \omega \left[ (E_l - E_n)^2 + (\hbar \Gamma_{ln} + i\hbar \omega)^2 \right]}, \tag{7.19}$$

143
Here \( r_{nm}^\pm = \langle n | \hat{r}_m^\pm | m \rangle \) and \( \hat{r}_m^\pm = (\hat{x}_m \pm i\hat{y}_m)/2 \). The above three terms have different physical meanings. The first term comes from interband effect and the latter two terms are related to intraband contribution. The tricky part is the orbital magnetization, which reads [244, 245]

\[
M_{\text{orb}} = M_{\text{orb}}^{LC} + M_{\text{orb}}^{IC} = \frac{e}{2\hbar c} \text{Im} \int \frac{d^3k}{(2\pi)^3} \langle \partial_k u_k | \times (H_k + E_k) | \partial_k u_k \rangle .
\]

(7.20)

Here we use tight-binding approach to investigate the inverse Faraday effect.

A similar effect, optical rectification, refers to a second order static electric polarization induced by a polarized light. The optical rectification susceptibility tensor is given by [246]

\[
\chi_{\text{rect}}^{abc} = B^{abc}(\omega) + B^{acb^*}(\omega) + \frac{1}{2} \partial_\omega \text{Re} \left( \Lambda^{abc}(\omega) \right),
\]

(7.21)

where

\[
\text{Im} \Lambda^{abc}(\omega) = -\frac{i\pi e^3}{2\hbar^2} \int \frac{dk}{8\pi^3} \sum_{nm\sigma} f_{nm} \left( r_{nm}^b (r_{nm}^c)_{k\sigma} + r_{mn}^c (r_{nm}^b)_{k\sigma} \right) \delta (\omega_{mn} - \omega),
\]

(7.22)

and

\[
\text{Im} B^{abc}(\omega) = \frac{\pi e^3}{2\hbar^2} \int \frac{dk}{8\pi^3} \sum_{nm\sigma} f_{nm} \left[ \frac{r_{nm}^a (r_{mn}^c)_{k\sigma} + r_{mn}^c (r_{nm}^a)_{k\sigma}}{\omega_{pm}} \right] \delta (\omega_{mn} - \omega) .
\]

(7.23)

The real part of the susceptibility can be obtained by using Kramers-Kronig relation. We have implemented these effects with both first-principles theory and tight-binding approach. More details will be found in future publications.

7.3 Nonlinear Magnetophotocurrent Responses

Recently, shift photocurrent and injection photocurrent have attracted a lot of interests in time-reversal symmetric system. Here we propose two new types of second-order nonlinear dc photocurrents in both time-reversal symmetry and inversion symmetry broken system, namely, magnetic shift photocurrent and magnetic injection photocurrent. These photocurrents can be detected
in asymmetric valley-splitting and noncentrosymmetric Chern insulator system. Magnetic shift current induced by circularly polarized light is dominated by shift vector and inter-Berry curvature, while magnetic injection current is determined by absorption strength and asymmetry of the group velocity at -\( k \) and \( k \) points. We showed the magnetic photocurrents are counterparts of shift and injection photocurrents. We separate the product of the electric field into symmetric and antisymmetric parts when exchanging the Cartesian indices \( b \) and \( c \)

\[
E^b(\omega)E^c(\omega) = \text{Re} \left( E^b E^c^* \right) + i \text{Im} \left( E^b E^c^* \right).
\]

(7.24)

After symmetrization with \( \pm \omega \), we have the injection current

\[
\frac{dJ^a}{dt} = -\frac{2\pi e^3}{\hbar^2 V} \sum_{b,c} \Delta_{bc} \left\{ r^b_{cv}, r^c_{ve} \right\} \text{Re} \left( E^b E^c^* \right) \delta (\omega_{cv} - \omega)
- i\frac{2\pi e^3}{\hbar^2 V} \sum_{b,c} \left[ r^b_{cv}, r^c_{ve} \right] \text{Im} \left( E^b E^c^* \right) \delta (\omega_{cv} - \omega).
\]

(7.25)

That is,

\[
\frac{d \langle J^a \rangle^{(2)}}{dt} = \sum_{b,c} \sigma_{2\text{MIC}}^{abc} \text{Re} \left( E^b E^c^* \right) + \sum_{b,c} i\eta_{2\text{MIC}}^{abc} \text{Im} \left( E^b E^c^* \right),
\]

(7.26)

where

\[
\sigma_{2\text{MIC}}^{abc} = \frac{e^3 \pi}{2\hbar^3} \int \frac{dk}{(2\pi)^3} \sum_{nm\sigma} f_{nm} \Delta_{nm}^a \left\{ r^b_{mn}, r^c_{nm} \right\} \delta (\omega_{mn} - \omega).
\]

(7.27)

The term \( \sigma_{2\text{MIC}}^{abc} \) is vanishing in time-reversal symmetric system but remains nonzero if -\( k \) and \( k \) band symmetry is broken. We name it as magnetic injection current (MIC). The second term is the known injection current tensor, given by

\[
\eta_{2\text{MIC}}^{abc} = \frac{e^3 \pi}{2\hbar^3} \int \frac{dk}{(2\pi)^3} \sum_{nm\sigma} f_{nm} \Delta_{nm}^a \left[ r^b_{mn}, r^c_{nm} \right] \delta (\omega_{mn} - \omega).
\]

(7.28)

Similarly, we have magnetic shift current (MSC)

\[
J_{\text{MSC}}^a = \frac{2\pi e^3}{\hbar^2} \int \frac{dk}{(2\pi)^3} f_{mn} \left( r_{mn;ka}^c r_{mn;ka}^b - r_{mn;ka}^b r_{mn;ka}^c \right) \delta (\omega_{mn} - \omega) \text{Im} \left( E^b(\omega) E^c(-\omega) \right).
\]

(7.29)
Under time-reversal symmetry, MSC can be rewritten by

\[
J_{\text{MSC}}^a = \frac{2\pi e^3}{\hbar^2} \int \frac{dk}{(2\pi)^d} \sum_{n,m,\sigma} f_{nm} R_{nm}^a(k) \Omega_{nm}^a(k) \delta(\omega_{nm} - \omega) \text{Im}(E_b E_c^*) .
\]

(7.30)

The MSC is vanishing since Berry curvature \( \Omega_{nm}^a \) is odd function and shift vector is even function with \( k \) under time reversal symmetry.

![Figure 7.4: Atomic structure of bilayer antiferromagnetic MnBi\(_2\)Te\(_4\). The magnetic orderings on Mn atoms are marked by blue (down) and red (up) arrows. The right part is the band structure and band energy difference in the Brillouin zone between VBM and CBM.](image)

Here we use bilayer antiferromagnetic (AFM) MnBi\(_2\)Te\(_4\) as an example. The structure of bilayer MnBi\(_2\)Te\(_4\) is shown in Figure 7.4. AFM MnBi\(_2\)Te\(_4\) has attracted great interest as it provides a new solid-state platform for exploring emerging magnetic topological physics from both experiment and theory. Mong et al. proposed new states of AFM topological insulator characterized by
a $Z_2$ topological invariant [247]. Remarkably, it shows the axion insulating state (first predicted as an elementary particle in extensions of the standard model in particle physics) with gapped bulk and surface states. AFM MnBi$_2$Te$_4$ covers two hot topics of quantum anomalous Hall effect and 2D ferromagnetism. The same system provides a great opportunity to investigate the NLO phenomena. As shown in Figure 7.4, the band energy symmetry at $k$ and $-k$ is broken, which enables the magnetic injection current under linearly polarized light irradiation.

We calculated the magnetic injection current response and found it is magnetically switchable between AFM and $\mathcal{T}$ AFM states. Time reversal operator $\mathcal{T}$ switches the magnetic orientation of Mn atoms on each layer. In fact, bilayer AFM MnBi$_2$Te$_4$ is $\mathcal{PT}$ symmetric, however, has neither inversion symmetry nor time reversal symmetry. However, SC vanishes in $\mathcal{PT}$ symmetric system, with $\mathcal{PT}$, we can take wave functions to be real (due to the transformation) and shift vector $\mathbf{R}$ and Berry curvature $\Omega$ will vanish in this case. To enable the SC and Berry curvature, we apply a perpendicular electric field to break the $\mathcal{PT}$ symmetry. More specifically, electric field $E_z$ breaks the mirror symmetry $\mathcal{M}_z$, hence, breaks $\mathcal{M}_y\mathcal{M}_z$. In AFM-$z$ case with $\mathcal{PT}$, $C_{3z}, \mathcal{M}_x\mathcal{T}$, and $\mathcal{M}_y\mathcal{M}_z$ symmetry, we have

$$J_{\text{MIC}}^{xxx} = -J_{\text{MIC}}^{xyy} = -J_{\text{MIC}}^{yxy} = -J_{\text{MIC}}^{yyx}. \quad (7.31)$$

The microscopic distribution of MIC in AFM and $\mathcal{T}$ AFM are shown in Figure 7.5 and Figure 7.6. It shows that the MIC is indeed magnetically switchable. The $J_{\text{MIC}}^{yyy}$ and $J_{\text{MIC}}^{yxx}$ are vanishing due to the $\mathcal{M}_y\mathcal{M}_z$ symmetry. We have the transformation relation

$$\mathcal{M}_y\mathcal{M}_z J_{\text{MIC}}^{yyy} = -J_{\text{MIC}}^{yyy} = J_{\text{MIC}}^{yyy}. \quad (7.32)$$

This is very different with the SC response under external perpendicular electric field $E_z$. SC is insensitive to time-reversal operator and involves the optical absorption, which essentially results in the dissipation. The effect of dissipation is crucial for nonreciprocal current response which is forbidden in linear response under time-reversal symmetry, corresponding to the macroscopic irreversibility. Hence, SC is a type of nonlinear and nonreciprocal optical responses in nontrosym-
metric systems.

Figure 7.5: The distribution of $\Delta x_{vc} \{ r_{cv}^x, r_{vc}^x \} \delta (\omega_{mn} - \omega)$ in Brillouin zone at $\omega=0.12$ eV and $\omega=0.60$ eV in AFM bilayer MnBi$_2$Ti$_4$.

Figure 7.6: The distribution of $\Delta x_{vc} \{ r_{cv}^x, r_{vc}^x \} \delta (\omega_{mn} - \omega)$ in Brillouin zone at $\omega=0.12$ eV and $\omega=0.60$ eV in AFM bilayer MnBi$_2$Ti$_4$. 
For SC under electric field $E_z$, we have

$$J_{yxx}^{\text{SC}} = -J_{yxy}^{\text{SC}} = -J_{xyx}^{\text{SC}} = -J_{xxy}^{\text{SC}}.$$  

(7.33)

It should be noted that $J_{xxx}^{\text{SC}}$ is always zero due to $\mathcal{M}_x \mathcal{T}$ symmetry. Figure 7.7 shows the large SC response and electrically switchable SC in bilayer AFM MnBi$_2$Te$_4$. The microscopic distribution of SC intensity in $k$ space is shown in Figure 7.8.

Figure 7.7: Shift current in bilayer AFM MnBi$_2$Te$_4$ under electric field $E_z = \pm 0.01 \text{V/Å}$ and convergence of $k$-point sampling in the tight-binding calculation

Figure 7.8: The distribution of $\sum_{nm\sigma} f_{nm} R_{nm}^\sigma(k) r_{nm}^\sigma r_{nm}^\delta (\omega_{mn} - \omega)$ in the first Brillouin zone at $\omega = 0.36$ eV and $\omega = 0.60$ eV in AFM bilayer MnBi$_2$Te$_4$. 

149
7.4 Nonlinear Current Response with External Magnetic Field

The semiclassical equations of motion of an electron wave packet in a metal are given by

$$\begin{align*}
\dot{r} &= \frac{1}{\hbar} \nabla_k \varepsilon_k - \hat{k} \times \Omega \\
\hbar \dot{k} &= -eE - e\dot{r} \times B \\
e &> 0 \text{ (charge convention)} \\
\varepsilon_k &= \varepsilon^0_k - \mathbf{M} \cdot \mathbf{B}
\end{align*}$$

(7.34)

Here, we consider intraband effect only. We then obtain

$$\begin{align*}
\dot{r} &= \frac{1}{\hbar} \nabla_k \varepsilon_k - \frac{1}{\hbar} (-eE - e\dot{r} \times \mathbf{B}) \times \Omega \\
&= \frac{1}{\hbar} \left( \nabla_k \varepsilon_k + e \mathbf{E} \times \Omega + e (\mathbf{r} \cdot \Omega) \mathbf{B} - e (\mathbf{B} \cdot \Omega) \dot{r} \right) \\
&= \frac{1}{\hbar} \left( \nabla_k \varepsilon_k + e \mathbf{E} \times \Omega + e \left( \left( \frac{1}{\hbar} \nabla_k \varepsilon_k - k \times \Omega \right) \cdot \Omega \right) \mathbf{B} - e (\mathbf{B} \cdot \Omega) \dot{r} \right)
\end{align*}$$

(7.35)

Finally, one arrives at

$$\dot{r} = \left( 1 + \frac{e}{\hbar} \mathbf{B} \cdot \Omega \right)^{-1} \frac{1}{\hbar} \left( \nabla_k \varepsilon_k + e \mathbf{E} \times \Omega + e \left( \frac{1}{\hbar} \nabla_k \varepsilon_k \cdot \Omega \right) \mathbf{B} \right),$$

(7.36)

and

$$\hbar \dot{k} = \left( 1 + \frac{e}{\hbar} \mathbf{B} \cdot \Omega \right)^{-1} \left( -eE - \frac{e}{\hbar} \nabla_k \varepsilon_k \times \mathbf{B} - \frac{e^2}{\hbar} (\mathbf{E} \cdot \mathbf{B}) \Omega \right)$$

(7.37)

The modified density of states reads (field induced volume change of the phase space) [245]

$$D_{nk} = \frac{\left( 1 + \frac{e}{\hbar} \mathbf{B} \cdot \Omega_{nk} \right)}{(2\pi)^d}.$$ 

(7.38)

The orbital magnetization can be derived from the semiclassical theory. The current density are given by

$$\begin{align*}
\mathbf{j} &= -e \int [dk] \left( \mathbf{v}_k + \frac{\varepsilon_k}{\hbar} \mathbf{E} \times \Omega + \frac{\varepsilon_k}{\hbar} (\mathbf{v}_k \cdot \Omega) \mathbf{B} \right) f, \\
\mathbf{v}_k &= v_k - \frac{1}{\hbar} \nabla_k (m \cdot B).
\end{align*}$$

(7.39)
To get the distribution function $f$, we apply Fourier expansion

$$f = f_0 + f_1 e^{-i\omega t} + f_2 e^{-2i\omega t} + f_3 e^{-3i\omega t} + \cdots + f_1^0 + \frac{f_2^0}{\tau} + \frac{f_3^0}{\tau^2} + \frac{f_4^0}{\tau^3} + \cdots$$  (7.40)

Using the Boltzmann transport equation

$$\frac{df}{dt} = \dot{k} \cdot \nabla_k f + \partial_t f,$$  (7.41)

we then obtain

$$\frac{1}{D} \left( -eE - \frac{e^2}{\hbar} (E \cdot B) \Omega \right) \cdot \nabla_p f_0 - i\omega f_1 = -\frac{f_1}{\tau},$$

$$\frac{1}{D} \left( -eE - \frac{e^2}{\hbar} (E \cdot B) \Omega \right) \cdot \nabla_p f_1 - 2i\omega f_2 = -\frac{f_2}{\tau},$$

$$\frac{1}{D} \left( -eE - \frac{e^2}{\hbar} (E \cdot B) \Omega \right) \cdot \nabla_p f_2 - 3i\omega f_3 = -\frac{f_3}{\tau}.\quad (7.42)$$

Hence the distribution functions are

$$f_1 = \frac{-\tau}{1 - i\omega \tau} \frac{1}{D} \left( -eE - \frac{e^2}{\hbar} (E \cdot B) \Omega \right) \nabla_p f_0,$$

$$f_2 = \frac{\tau^2}{(1 - i\omega \tau)(1 - 2i\omega \tau)} \frac{1}{D^2} \left[ \left( -eE - \frac{e^2}{\hbar} (E \cdot B) \Omega \right) \nabla_p \right]^2 f_0,$$  (7.43)

$$f_3 = \frac{\tau^2}{(1 - i\omega \tau)(1 - 2i\omega \tau)(1 - 3i\omega \tau)} \frac{1}{D^3} \left[ \left( -eE - \frac{e^2}{\hbar} (E \cdot B) \Omega \right) \nabla_p \right]^3 f_0.$$

Numerous magneto-optical phenomena can be readily distinguished from the above equations. More specifically, the metals with large Berry curvature (e.g. topological Dirac/Weyl semimetals) are nice platforms for exploring fundamental nonlinear magneto-optical responses.

### 7.5 Higher Order Nonlinear Optical Effects and Many-Body Effects

Two-photon absorption (TPA) is a third order nonlinear optical response in which the media absorbs two photons simultaneously. One can obtain two-photon absorption rate from the Fermi’s
We then obtain the two-photon absorption coefficient as follows:

\[
\Gamma_{vck} = \frac{2\pi}{\hbar} \left( \frac{eA_1}{mc} \right)^2 \left( \frac{eA_2}{mc} \right)^2 \sum_m \frac{\langle c| e_2 \cdot p|m \rangle \langle m| e_1 \cdot p|v \rangle}{E_m - E_v - \hbar\omega_1} + \frac{\langle c| e_2 \cdot p|m \rangle \langle m| e_1 \cdot p|v \rangle}{E_m - E_v - \hbar\omega_2} \left| \delta (E_c(k) - E_v(k) - \hbar\omega_1 - \hbar\omega_2) \right| \tag{7.44}
\]

We then obtain the two-photon absorption coefficient as follows:

\[
\alpha(\omega) = \frac{8\pi^3\hbar^2 N^2}{cm^4 n_1 n_2^2 \omega_1 \omega_2} \int_{BZ} \frac{2dk}{(2\pi)^3} \left| \sum_m \frac{\langle c| e_2 \cdot p|m \rangle \langle m| e_1 \cdot p|v \rangle}{E_m - E_v - \hbar\omega_1} + \frac{\langle c| e_2 \cdot p|m \rangle \langle m| e_1 \cdot p|v \rangle}{E_m - E_v - \hbar\omega_2} \delta (E_c(k) - E_v(k) - \hbar\omega_1 - \hbar\omega_2) \right| \tag{7.45}
\]

Similarly, we consider the indirect band-to-band transition with phonon contributions (electron-phonon coupling, denoted by \( g \)). The transition rate is given by

\[
\Gamma_{v_{k1}\rightarrow v_{k2}} = \frac{2\pi}{\hbar} \left( \frac{eA_0}{mc} \right)^2 \left| \sum_m \frac{\langle c| g|m \rangle n_q^{1/2} \langle m| e \cdot p|v \rangle}{E_m(k_1) - E_v(k_1) - \hbar\omega} \right|^2 + \frac{\langle c| g|m \rangle (n_q + 1)^{1/2} \langle m| e \cdot p|v \rangle}{E_m(k_2) - E_v(k_1) \pm \hbar\omega_q} \left| \delta (E_c(k_2) - E_v(k_1) - \hbar\omega \pm \hbar\omega_q) \right| \tag{7.46}
\]

where the momentum conversion gives \( q = k_2 - k_1 \). The corresponding absorption coefficient is given by

\[
\alpha(\omega) = \frac{4\pi^2 \hbar^2}{ncm^2\omega} \int_{BZ} \int_{BZ} \frac{2dk_2dk_1}{(2\pi)^3} \left| \sum_m \frac{\langle c| g|m \rangle (n_q + \frac{1}{2} \pm \frac{1}{2})^{1/2} \langle m| e \cdot p|v \rangle}{E_m(k_1) - E_v(k_2) \pm \hbar\omega_q} \right|^2 + \frac{\langle c| g|m \rangle (n_q + \frac{1}{2} \pm \frac{1}{2})^{1/2} \langle m| e \cdot p|v \rangle}{E_m(k_2) - E_v(k_1) \pm \hbar\omega_q} \left| \delta (E_c(k_2) - E_v(k_1) - \hbar\omega \pm \hbar\omega_q) \right| \tag{7.47}
\]

The upper (lower) sign denotes phonon emission (absorption) process.

Two-particle Bethe-Salpeter equation (BSE) can be solved for pair excitations which considers excitonic and local-field contributions to both linear and nonlinear optical responses. Using the electron-hole pair states, Chan et al. [248] employed a first-principles interacting Green’s-function
on the Keldysh contour to study nonlinear photocurrents. They start with the equation of motion for the interacting single-particle Green’s function $G$ on the Keldysh contour $C$ [248]

$$\left[ i \frac{d}{dt} - H(t) \right] G(t, t') = \delta(t, t') \int_C \Sigma(t, t'') G(t, t'') \, dt'',$$

$$i \hbar \frac{\partial}{\partial t} \rho_{nm,k}(t) = \left[ H_{0,k}(t) + \Sigma^G_{k} + \delta \Sigma_k(t) - e E \cdot r, \rho \right].$$

(7.48)

The approximated SC susceptibility for a simple two-band model with exciton states labeled by $m$ and $n$ is given by [248]

$$\sigma_{abc}(0; \omega, -\omega) \sim -\frac{e^3}{8 \pi^2 m} \sum_{mn} \frac{P_m Q_n R^c_{mn}}{E_m (E_n - \hbar \omega - i \eta)},$$

$$P_m = \sum_k \phi^m \phi_{vck} \phi_{ck},$$

$$Q_{mn} = i \langle m | \partial_k | n \rangle,$$

$$R_m = \sum_k \phi^m \phi_{ck} \phi^{\text{inter}}_{vck} = \langle 0 | r^{\text{inter}} | n \rangle.$$  

(7.49)

Interestingly, they found bound exciton states can generate significant SC within the band gap.

All the optical responses we have discussed so far are limited in dipole approximation. However, multipole radiation may also contribute in some cases. Basically, we can expand the exponential in the matrix element in a power series

$$e^{ik \cdot r} = 1 + ik \cdot r - \frac{(k \cdot r)^2}{2!} + \cdots,$$

(7.50)

and we obtain

$$\langle f | e^{ik \cdot r} \cdot \mathbf{p} | i \rangle = \langle f | \mathbf{p} | i \rangle + \langle f | ik \cdot r \mathbf{p} | i \rangle - \left( f \left| \frac{(k \cdot r)^2}{2!} \mathbf{p} \right| i \right) + \cdots,$$

(7.51)

If we take the $y$-axis along $e$ and the $x$-axis along $k$, we can transform the second term to the form
with electric quadrupole and magnetic dipole contribution

\[
M_z = -\hbar k \left\{ f \left| x \frac{\partial}{\partial y} i \right| + \frac{i}{2} \hbar k \left\{ \left( f \left| x \frac{\partial}{\partial y} + y \frac{\partial}{\partial x} i \right| \right) + \left( f \left| x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} i \right| \right) \right\}. \tag{7.52}
\]

Using the definition of angular momentum,

\[
L_z = -i \hbar \left( f \left| y \frac{\partial}{\partial x} i \right| \right). \tag{7.53}
\]

we finally obtain the multipole terms

\[
M_z = -\frac{1}{2} i m k \omega_{ij} \langle f \left| xy \right| i \rangle + \frac{1}{2} k \langle f \left| L_z \right| i \rangle. \tag{7.54}
\]

The above discussions and NLO effects present the exciting new directions and important issues which will be addressed in future.

### 7.6 Conclusions

To summarize, a theoretical framework and computational approach have been developed for the discovery and design of nonlinear optical materials. We discovered the extraordinary SHG, shift current, and circular current and their switching behavior in multiferroic group IV monochalcogenide monolayers. The ferroicity-driven nonlinear SHG and photocurrent switching are unique to multiferroics where the direction of second order NLO responses is strongly correlated with intrinsic ferroic orders and external light polarization. Furthermore, we have presented an underexplored realm beyond classical linear Hall effect and conventional ferroelectrics with appealing extra opportunities for FNAHE-based nonlinear quantum electronics using ferroelectric and topological semimetals, e.g., few-layer WTe\(_2\). The results presented demonstrate that the fundamental understanding of the nonlinear optical phenomena together with their symmetry principles may offer unprecedented opportunities for the discovery and design of nonlinear optical materials and enable novel devices such as nonlinear quantum electronics, spintronics, magneto-electronics, and dynamic quantum materials which may promote the second quantum evolution.
REFERENCES


A.1 The Representation of Position Operator

The representation of position operator is not straightforward in the case that the operator $r$ is unbounded and does not have the complete Hilbert space for its domain. We use the wavefunction in the linear combination of Bloch functions

$$\varphi(r) = \frac{1}{(2\pi)^3} \int d^3k e^{i k \cdot r} \sum_n u_{nk}(r)a_n(k). \quad (A.1)$$

Then we operate $r = i\partial_k$ on the wavefunction

$$r\varphi(r) = \frac{1}{(2\pi)^3} \int d^3k e^{i k \cdot r} \sum_n r u_{nk}(r)a_n(k)$$

$$= \frac{1}{(2\pi)^3} \int d^3k \sum_n \psi_{nk}(r)a_n(k)(-i\partial_k) + \frac{1}{(2\pi)^3} \int d^3k \sum_n (i\partial_k)(u_{nk}(r)a_n(k)) \quad (A.2)$$

$$= \frac{1}{(2\pi)^3} \int d^3k \sum_n \psi_{nk}(i\partial_k a_n + \sum_m \xi_{nm} a_m).$$

We can then obtain the position matrix elements in the thermal dynamic limit

$$\langle nk|mk' \rangle = \delta(k - k') \xi_{nm} + i\delta_{nm} \nabla_k \delta(k - k'). \quad (A.3)$$

The physical meaning of the position operator is clear in the Wannier representation.

A.2 Derivation of Commutative Relations

The covariant derivative of an operator matrix elements is defined as

$$O_{nm;a} = \left[ \frac{\partial}{\partial k_a} - i \left( \xi_{nm}^a - \xi_{nm}^a \right) \right] O_{nm}. \quad (A.4)$$
The derivation of the commutative relation between intra position operator and other simple operator is shown as follows. First, we have

\[
\langle nk | [r_i, O] | mk' \rangle = \langle nk | r_i O - O r_i | mk' \rangle \\
= \sum_{l, k'} \langle nk | r_i | lk'' \rangle \langle lk'' | O | mk' \rangle - \sum_{l, k'} \langle nk | O | lk'' \rangle \langle lk'' | r_i | mk' \rangle
\]  
(A.5)

One rewrites the first term

\[
\sum_{i, k'} \langle nk | r_i | lk'' \rangle \langle lk'' | O | mk' \rangle \\
= \sum_{l, k''} \delta_{nl} \left[ \delta(k - k'') \xi_{nn} + i \nabla_k \delta(k - k'') \right] \delta(k' - k'') O_{lm}(k') \\
= \sum_{k''} \left[ \delta(k - k'') \xi_{nn} + i \nabla_k \delta(k - k'') \right] \delta(k' - k'') O_{nm}(k') \\
= \sum_{k''} \delta \left( k - k'' \right) \left[ \nabla_k \delta \left( k - k'' \right) \xi_{nm}(k') + i \sum_{k''} \left[ \nabla_k \delta \left( k - k'' \right) \right] \delta \left( k' - k'' \right) O_{nm}(k') \right] \\
= \delta \left( k - k' \right) \xi_{nm} O_{nm}(k') + i \nabla_k \left[ \delta(k - k') O_{nm}(k') \right].
\]  
(A.6)

Similarly, we can obtain the second term

\[
\sum_{l, k''} \langle nk | O | lk'' \rangle \langle lk'' | r_i | mk' \rangle \\
= \sum_{l, k''} \delta \left( k - k'' \right) O_{nl}(k'') \delta_{lm} \left[ \delta \left( k' - k'' \right) \xi_{nn} + i \nabla_k \delta \left( k' - k'' \right) \right] \\
= \sum_{k''} \delta(k - k'') O_{nm}(k') \left[ \delta(k' - k'') \xi_{mm} + i \nabla_{k''} \delta(k' - k'') \right] \\
= \delta \left( k - k' \right) \xi_{mm} O_{nm}(k') + \sum_{k''} \delta \left( k - k'' \right) O_{nm}(k') \left[ i \nabla_{k''} \delta \left( k' - k'' \right) \right] \\
= \delta \left( k - k' \right) \xi_{mm} O_{nm}(k') + O_{nm}(k) \left[ i \nabla_k \delta \left( k' - k \right) \right].
\]  
(A.7)

We also have the identity

\[
\delta(k - k') O_{nm}(k') = \delta(k' - k) O_{nm}(k). 
\]  
(A.8)
Finally, we can obtain the commutative relation by substituting the above equations

\[ \langle n k | r_z, O \rangle m' k' = i \delta (k - k') \nabla_k O_{nm}(k) + i \delta (k - k') O_{nm} (\xi_{nm} - \xi_{mn}). \]  

(A.9)

A.3 Levi-Civita Symbol and Cross Product

The Levi-Civita symbol \( \epsilon_{ijk} \) is a rank-3 tensor, defined by

\[
\epsilon_{ijk} = \begin{cases} 
0 & \text{if any two labels are the same} \\
1 & \text{if } i, j, k \text{ is an even permutation of } 1, 2, 3 \\
-1 & \text{if } i, j, k \text{ is an odd permutation of } 1, 2, 3 
\end{cases}
\]

(A.10)

The cross product of two vectors can be written by a determinant

\[
\mathbf{a} \times \mathbf{b} = \det \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \epsilon_{ijk} \hat{e}_i a_j b_k,
\]

(A.11)

where Einstein summation is applied to simplify the expressions. We therefore have the cross product formula in terms of Levi-Civita symbol

\[
(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k.
\]

(A.12)

The product of two Levi-Civita symbols can be simplified as

\[
\epsilon_{ijk} \epsilon_{imn} = \delta_{ii} (\delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}) + \delta_{im} \delta_{jn} \delta_{ki} + \delta_{in} \delta_{ji} \delta_{km} - \delta_{im} \delta_{ji} \delta_{kn} - \delta_{in} \delta_{jm} \delta_{ki} \\
= 3 (\delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}) + \delta_{km} \delta_{jn} + \delta_{jn} \delta_{km} - \delta_{jm} \delta_{kn} - \delta_{kn} \delta_{jm}
\]

(A.13)

\[= \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}.\]
Hence we can demonstrate the identity

\[ a \times (b \times c) = b(a \cdot c) - c(a \cdot b), \]  

(A.14)

since

\[ a \times (b \times c)_m = \epsilon_{mn}a_n (\epsilon_{ijk}b_j c_k) \]

\[ = \epsilon_{mn} \epsilon_{ijk} a_n b_j c_k \]

\[ = (\delta_{mj} \delta_{nk} - \delta_{mk} \delta_{nj}) a_n b_j c_k \]  

(A.15)

\[ = b_m a_k c_k - c_m a_j b_j \]

\[ = [b(a \cdot c)]_m - [c(a \cdot b)]_m \]

A.4 Covariant Derivative and Sum Rules

Considering the commutation relation

\[ [r^a,b^b] = [r^a_i,b^b] + [r^a_e,b^b] = i\hbar \delta_{ab}, \]

(A.16)

we can find

\[ \langle nk | [r^a,b^b] | mk \rangle = i\hbar \delta_{ab} \delta_{nm} = \langle nk | [r^a_i,b^b] | mk \rangle + \langle nk | [r^a_e,b^b] | mk \rangle \]

\[ = i \langle p^b_{nm} | k_a \rangle + \langle nk | [r^a \cdot p^b] | mk \rangle = i \langle p^b_{nm} | k_a \rangle + \langle nk | r^a \cdot p^b - p^b r^a | mk \rangle \]  

(A.17)

\[ = i \langle p^b_{nm} | k_a \rangle + \sum_{l \neq n} \langle nk | r^a_{l} | l k \rangle \langle l k | p^b | m k \rangle - \sum_{l \neq m} \langle nk | p^b_{l} | l k \rangle \langle l k | r^a_{l} | m k \rangle, \]

therefore we have

\[ i \langle p^b_{nm} | k_a \rangle = i\hbar \delta_{ab} \delta_{nm} - \left( \sum_{l \neq n} (r^a_{nl} | p^b_{lm} \rangle) - \sum_{l \neq m} (p^b_{nl} r^a_{lm}) \right). \]

(A.18)

The sum rule can be obtained from the identity

\[ \langle n | H | m \rangle = E_n \delta_{nm}. \]

(A.19)
Taking the first derivative, one can reach
\[
\partial_a (n|H|m) = (\partial_a n|H|m) + (n|\partial_a H|m) + (n|H|\partial_a m) = \partial_a E_m \delta_{nm}
\]
\[
= i E_m r_{nm}^a - i E_n r_{nm}^a + (n|\partial_a H|m) \tag{A.20}
\]
\[
= i \hbar \omega_{mn} r_{nm}^a + (n|\partial_a H|m).
\]

Taking the second derivative, one can obtain
\[
\partial_b \partial_a (n|H|m) = \partial_b \left( i \hbar \omega_{mn} r_{nm}^a + (n|\partial_a H|m) \right) \\
= i \hbar \Delta_{mn}^b r_{nm}^a + i \hbar \omega_{mn} \partial_b (r_{nm}^a) + (\partial_b n|\partial_a H|m) + (n|\partial_b \partial_a H|m) \tag{A.21}
\]
\[
+ (n|\partial_a H|\partial_b m),
\]
and we have the identity
\[
(\partial_a n|\partial_b m) = \sum_l (\partial_a n|l) (l|\partial_b m) = \sum_l r_{nl}^a r_{lm}^b. \tag{A.22}
\]

We first simplify the third term in the second derivative
\[
(\partial_b n|\partial_a H|m) = \left( \partial_b n \left| \partial_a \left( \sum_m |m\rangle E_m (m) \right) \right| m \right) \\
= \left( \partial_b n \left| \sum_l |\partial a l\rangle E_l (l) \right| m \right) + (\partial_b n |(\sum_l |l\rangle \partial a E_l (l)| m) \\
+ \left( \partial_b n \left| \sum_l |l\rangle E_l (l) \right| m \right) \tag{A.23}
\]
\[
= \sum_l (-r_{nl}^b E_l r_{lm}^a) + i \partial_a E_m r_{nm}^b + E_m (\partial_b n|\partial_a m) \\
= \hbar \sum_l (r_{nl}^b E_m r_{lm}^a \omega_{ml}) + i \partial_a E_m r_{nm}^b.
\]
The last term in the second derivative can be simplified as follows,

\[
\langle n | \partial_a H \partial_b m \rangle = \left( n \left| \partial_a \left( \sum_m |m\rangle E_m(m) \right) \right| \partial_b m \right)
\]
\[
= \left( n \sum_l |\partial_a l\rangle E_l(l) \right| \partial_b m \rangle \right) + \left( n |(\Sigma_l|l\rangle \partial_a E_l(l)| \partial_b m \rangle \right)
\]
\[
= \sum_l \left( -r_{nl}^a E_l r_{lm}^b \right) - i\partial_a E_n r_{nm}^b + E_n (\partial_a n|\partial_b m) \]
\[
= \hbar \sum_l \left( r_{nl}^a r_{lm}^b \omega_{nl} \right) - i\partial_a E_n r_{nm}^b. \tag{A.24}
\]

We then arrive at

\[
i \Delta_m^b r_{nm}^a + i\omega_{nm} \partial_b (r_{nm}^a) + \sum_l \left( r_{nl}^a r_{lm}^b \omega_{nl} + r_{nl}^b r_{lm}^a \omega_{ml} \right) + i \Delta_m^a r_{nm}^b + \omega_{nm}^{ab} = 0. \tag{A.25}
\]

Finally, we obtain the following sum rule

\[
(r_{nm}^a)_b = \frac{1}{i\omega_{nm}} \left( -i \left( r_{nm}^a \Delta_m^b + r_{nm}^b \Delta_m^a \right) - \omega_{nm}^{ab} - \sum_{l \neq n, l \neq m} \left( \omega_{nl} r_{nl}^a r_{lm}^b - \omega_{lm} r_{nl}^b r_{lm}^a \right) \right)
\]
\[
= \frac{i}{\omega_{nm}} \left( i \left( r_{nm}^a \Delta_m^b + r_{nm}^b \Delta_m^a \right) + \omega_{nm}^{ab} + \sum_{l \neq n, l \neq m} \left( \omega_{nl} r_{nl}^a r_{lm}^b - \omega_{lm} r_{nl}^b r_{lm}^a \right) \right) \tag{A.26}
\]
\[
= \frac{i}{\omega_{nm}} \left( i \left( r_{nm}^a \Delta_m^b + r_{nm}^b \Delta_m^a \right) - \omega_{nm}^{ab} - \sum_{l \neq n, l \neq m} \left( \omega_{nl} r_{nl}^a r_{lm}^b - \omega_{lm} r_{nl}^b r_{lm}^a \right) \right).
\]