Electronic, Optical, and Thermal Properties of Reduced-Dimensional Semiconductors

Shouting Huang
Washington University in St. Louis

Follow this and additional works at: http://openscholarship.wustl.edu/etd

Recommended Citation
Huang, Shouting, "Electronic, Optical, and Thermal Properties of Reduced-Dimensional Semiconductors" (2013). All Theses and Dissertations (ETDs). 1137.
http://openscholarship.wustl.edu/etd/1137

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Electronic, Optical, and Thermal Properties of Reduced-Dimensional Semiconductors

by

Shouting Huang

A dissertation presented to the Graduate School of Arts and Sciences of Washington University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

August 2013
St. Louis, Missouri
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 THEORETICAL METHODS</td>
<td>21</td>
</tr>
<tr>
<td>2.1 Density Functional Theory (DFT)</td>
<td>22</td>
</tr>
<tr>
<td>2.1.1 Hohenberg-Kohn Theorem</td>
<td>22</td>
</tr>
<tr>
<td>2.1.2 Kohn-Sham Equations</td>
<td>25</td>
</tr>
<tr>
<td>2.1.3 Pseudopotential Method</td>
<td>31</td>
</tr>
<tr>
<td>2.1.4 Momentum-Space Formalism</td>
<td>37</td>
</tr>
<tr>
<td>2.2 Many-Body Perturbation Theory (MBPT)</td>
<td>40</td>
</tr>
<tr>
<td>2.2.1 GW Approximation</td>
<td>42</td>
</tr>
<tr>
<td>2.2.2 Bethe-Salpeter Equation (BSE) and Optical Absorption</td>
<td>53</td>
</tr>
<tr>
<td>2.3 First-Principles Phonon Calculations</td>
<td>63</td>
</tr>
<tr>
<td>2.3.1 Lattice Vibrational Modes in Harmonic Crystals</td>
<td>64</td>
</tr>
<tr>
<td>2.3.2 Linear Response Theory</td>
<td>70</td>
</tr>
</tbody>
</table>
### 3 BAND OFFSETS ENGINEERING IN SI/GE CORE-SHELL NANOWIRES  
#### 3.1 Introduction of Heterojunctions and Core-Shell Nanowires  
#### 3.2 First-Principles Simulation Setup  
#### 3.3 Electronic Structures in Si/Ge Core-Shell NWs  
#### 3.4 Strain Control of Band Offsets in Si/Ge Core-Shell NWs  
#### 3.5 Geometry Control of Band Offsets in Si/Ge Core-Shell NWs  
#### 3.6 Summary of this Chapter  

### 4 LATTICE VIBRATIONS IN SI/GE CORE-SHELL NANOWIRES  
#### 4.1 Introduction  
#### 4.2 First-Principles Simulation Setup  
#### 4.3 Vibrational Modes in Si/Ge Core-Shell NWs  
##### 4.3.1 High-Frequency Optical Modes  
##### 4.3.2 Transverse and Longitudinal Optical Modes Splitting  
##### 4.3.3 Raman Scattering Spectrum  
##### 4.3.4 Radial Breathing Modes  
#### 4.4 Summary of this Chapter  

### 5 EXCITON SPECTRA OF TWO-DIMENSIONAL SEMICONDUCTING CARBON STRUCTURES  
#### 5.1 Introduction to Excitons and Exciton Spectra  

---

iii
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>First-Principles Simulation Setup</td>
<td>119</td>
</tr>
<tr>
<td>5.3</td>
<td>Electronic Structures and Optical Properties of Graphyne</td>
<td>120</td>
</tr>
<tr>
<td>5.4</td>
<td>Exciton Spectra of Graphyne</td>
<td>124</td>
</tr>
<tr>
<td>5.5</td>
<td>Modified Hydrogenic Model for Excitons in 2D Semiconductors</td>
<td>128</td>
</tr>
<tr>
<td>5.6</td>
<td>Summary of this Chapter</td>
<td>136</td>
</tr>
<tr>
<td>6</td>
<td>ELECTRONIC STRUCTURES AND QUASIPARTICLE ENERGY OF SILICENE STRUCTURES</td>
<td>139</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction of Silicene</td>
<td>139</td>
</tr>
<tr>
<td>6.2</td>
<td>First-Principles Simulation Setup</td>
<td>142</td>
</tr>
<tr>
<td>6.3</td>
<td>Electronic Structures of Planar and Simply Buckled Silicene Structures</td>
<td>143</td>
</tr>
<tr>
<td>6.4</td>
<td>A Brief Review of Recently Proposed Silicene Structures</td>
<td>146</td>
</tr>
<tr>
<td>6.5</td>
<td>Electronic Structures of Experimentally Proposed Silicene Structures</td>
<td>151</td>
</tr>
<tr>
<td>6.6</td>
<td>Summary of this Chapter</td>
<td>162</td>
</tr>
<tr>
<td>7</td>
<td>CONCLUSIONS</td>
<td>163</td>
</tr>
<tr>
<td>7.1</td>
<td>Conclusions of Projects</td>
<td>163</td>
</tr>
<tr>
<td>7.2</td>
<td>Future Work</td>
<td>165</td>
</tr>
</tbody>
</table>

REFERENCES ............................... 167

VITA ....................................... 177
Working on the Ph.D. has been an unforgettable experience in my life. It’s my pleasure to recall the past years and express my acknowledgements to those who have given me support and encouragement along the way.

First and foremost I am deeply grateful to my advisor, Dr. Li Yang. It’s my honor to be his first Ph.D. student. His enormous encouragement and support brought me to the current milestone. I’m greatly impressed by his diligent working, thorough thinking, and precise attitude, which I believe shall also be valuable for my own career development.

I am greatly grateful to Dr. Zohar Nussinov for advising me in the early years of my graduate study. His instructions helped me in passing the qualifying exam. He has a brilliant mind and he is a really very nice person.

I am also grateful to the members of Dr. Yang’s group, especially Yufeng Liang and Ryan Soklaski. As my most close friend now, Yufeng is an easy-going and smart person. Over the past few years, we have had a comfortable experience in sharing an apartment. I can still remember his consideration when I was sick or when I went back late. Ryan is a very nice person, one of nicest American student that I have
ever met. He has devoted a lot of time in proofreading my manuscripts and even this thesis. It’s my pleasure to work with them in the same group.

Special thanks to Dr. Patrick Gibbons, Dr. Alexander Seidel, Dr. Cynthia Lo, and Dr. Richard Loomis for serving in my dissertation defense committee. I’m grateful to their instructions and support. And also thanks to many professors including but not be limited to Dr. Carl Bender, Dr. Mark Alford, Dr. Michael Ogilvie, Dr. Jr-Shin Li, Dr. Heinz Schaeftler, Dr. Guofu Zhou, Dr. Haojun Liang, and Dr. Xuejie Huang for their help.

I also thank many friends for providing support and friendship, and bringing me a colorful life. There are too many that I couldn’t completely list here. They includes Yunfei Shi, Chi Zhang, Xiaoqi Wang, Wei Wang, Zhenyu Zhou, Dandan Hu, Dongsu Du, Lingzhi Hu, Peng Yang, Bo Peng, Dong Ding, Bo Sun, Keyong Han, Yuechen Shen, Jing Li, Jiajing Xu, Jianli Pan, Lin Ma, Chengjie Wu, Peng Li, Jun Han, Ruixiang Fei, Zhi Shuang, Yong Wu, Zhiwen Shi, Chengping Wu, etc.

I gratefully acknowledge the funding sources and the computational resources that made my Ph.D. work possible. The funding sources include Washington University Fellowships and National Science Foundation. The computational resources have been provided by Lonestar of Teragrid at the Texas Advanced Computing Center and Hoppers of the National Energy Research Scientific Computing Center (NERSC).

Finally, and most importantly, I would like extend my great gratitude to my parents and my sisters. Their constant love and support always warm my heart and inspire my strength.
ABSTRACT

Electronic, Optical, and Thermal Properties of Reduced-Dimensional Semiconductors

by

Shouting Huang

Doctor of Philosophy in Physics,
Washington University in St. Louis, August, 2013.

Professor Li Yang, Chair

Reduced-dimensional materials have attracted tremendous attention because of their new physics and exotic properties, which are of great interests for fundamental science. More importantly, the manipulation and engineering of matter on an atomic scale yield promising applications for many fields including nanoelectronics, nanobiotechnology, environments, and renewable energy. Because of the unusual quantum confinement and enhanced surface effect of reduced-dimensional materials, traditional empirical models suffer from necessary but unreliable parameters extracted from previously-studied bulk materials. In this sense, quantitative, parameter-free approaches are highly useful for understanding properties of reduced-dimensional materials and, furthermore, predicting their novel applications.
The first-principles density functional theory (DFT) is proven to be a reliable and convenient tool. In particular, recent progress in many-body perturbation theory (MBPT) makes it possible to calculate excited-state properties, e.g., quasiparticle (QP) band gap and optical excitations, by the first-principles approach based on DFT. Therefore, during my PhD study, I employed first-principles calculations based on DFT and MBPT to systematically study fundamental properties of typical reduced-dimensional semiconductors, i.e., the electronic structure, phonons, and optical excitations of core-shell nanowires (NWs) and graphene-like two-dimensional (2D) structures of current interests.

First, I present first-principles studies on how to engineer band alignments of nano-sized radial heterojunctions, Si/Ge core-shell NWs. Our calculation reveals that band offsets in these one-dimensional (1D) nanostructures can be tailored by applying axial strain or varying core-shell sizes. In particular, the valence band offset can be efficiently tuned across a wide range and even be diminished via applied strain. Two mechanisms contribute to this tuning of band offsets. Furthermore, varying the size of Si/Ge core-shell NWs and corresponding quantum confinement is shown to be efficient for modifying both valence and conduction band offsets simultaneously. Our proposed approaches to control band offsets in nano-sized heterojunctions may be of practical interest for nanoelectronic and photovoltaic applications.

Additionally, I also studied the lattice vibrational modes of Si/Ge core-shell NWs. Our calculations show that the internal strain induced by the lattice mismatch between core and shell plays an important role in significantly shifting the frequen-
cy of characteristic optical modes of core-shell NWs. In particular, our simulation demonstrates that these frequency shifts can be detected by Raman-scattering experiments, giving rise to a convenient and nondestructive way to obtain structural information of core-shell materials. Meanwhile, another type of collective modes, the radial breathing modes (RBM), is identified in Si-core/Ge-shell NWs and their frequency dependence is explained by an elastic media model. Our studied vibrational modes and their frequency evolution are useful for thermoelectric applications based on core-shell nanostructures.

Then I studied optical properties and exciton spectra of 2D semiconducting carbon structures. The energy spectra and wavefunctions of excitons in the 2D graphene derivatives, \( i.e. \), graphyne and graphane, are found to be strongly modified by quantum confinement, making them qualitatively different from the usual Rydberg series. However, their parity and optical selection rules are preserved. Thus a one-parameter hydrogenic model is applied to quantitatively explain the \textit{ab initio} exciton spectra, and allows one to extrapolate the electron-hole binding energy from optical spectroscopies of 2D semiconductors without costly simulations. Meanwhile, our calculated optical absorption spectrum and enhanced spin singlet-triplet splitting project graphyne, an allotrope of graphene, as a good candidate for intriguing energy and biomedical applications.

Lastly, we report first-principles results on electronic structures of 2D graphene-like system, \( i.e. \), silicene. For planar and simply buckled silicene structures, we confirm their zero-gap nature and show a significant renormalization of their Fermi velocity
by including many-electron effects. However, the other two recently proposed silicene structures exhibit a finite band gap, indicating that they are gapped semiconductors instead of expected Dirac-fermion semimetals. This finite band gap of the latter two structures is preserved even with the Ag substrate included. The gap opening is explained by the symmetry breaking of the buckled structures. Moreover, our GW calculation reveals enhanced many-electron effects in these 2D structures. Finally the band gap of the latter two structures can be tuned in a wide range by applying strain.
1. INTRODUCTION

In modern physics, the fundamental approach to understand properties of solids is to study quantum mechanics including many-body interactions of electrons and nuclei. Thus, the structure and many other properties of solids rely on a good description of many-body interactions, particularly those many-electron interactions. It is, however, extremely challenging to directly employ the relevant Schrödinger equations to study realistic materials. Because of the large number of atoms in realistically-sized sample, the Schrödinger equations will have a huge number of variables, making computations formidable. More importantly, there is not an exact description of many-particle interactions that can simply be included in the corresponding Schrödinger equations. All these difficulties greatly obstruct a precise understanding of the fundamental properties of solids. Therefore, many theorems and approximations have been developed and applied in order to make the study feasible.

As a first step, the huge mass difference between electrons and ions makes it possible to employ the Born-Oppenheimer (BO) approximation that allows the wavefunction of materials to be broken into their electronic and nuclear components. Then the electronic wavefunctions are solved for while the nuclei can be fixed in a certain configuration, very often the equilibrium configuration. Focusing on the electronic
structure greatly simplifies the complexity of solid-state problems and it often provides good enough accuracy for studying most interesting properties. Meanwhile, when studying the nuclear motion, such as lattice vibrations, we can treat the electronic states as quick variables that always relax themselves instantaneously according the atomistic configurations, which substantially simplifies the calculation of the motions of nuclei.

We then focus on the electronic-structure calculations. For crystal structures which have spacial periodicity and translational symmetry, the Bloch theorem makes it possible to calculate the electronic states within the smallest unit cell commensurate with the period boundary condition and subsequently obtain all quantum states by adding an extra phase modulation. This theorem dramatically reduces the number of particles or the degrees of freedom involved in our study.

The single-particle picture of electrons has to be developed. In principle, each electron of solids cannot be described by a wavefunction determined by the single-particle Schrödinger equation that is independent of all the other electrons. However, for many studies, the independent-electron approximation is still a good way to simplify the actual problem of many-interacting electrons. In fact, the independent-electron approximation does not entirely neglect electron-electron (e-e) interactions. Rather, it assumes that most of the e-e effects can be taken into account with a treatment for the periodic effective potential appearing in the one-electron Schrödinger equation. Thus the effective potential includes not only the periodic potential due to the ions alone but also periodic effects due to all other e-e interactions. Of course, the
latter interaction depends on the configuration of the other electrons, which are also determined by a Schrödinger equation with the similar form.

One success of the single-particle description of electrons is the band theory. For a given wave vector and effective potential, there are a number of single-particle solutions, indexed by $n$, to the Schrödinger equation. These solutions, called bands, are separated in energy by a finite spacing at each $k$; if the band separation between occupied states and unoccupied states extends over all wave vectors, it is called a band gap. The band structure, i.e., the single-particle energy-momentum dispersion, is the collection of energy eigenstates within the first Brillouin zone (BZ). Many properties of electrons in a periodic potential can be calculated from their band structure and the associated wavefunctions [1, 2].

Although the independent-electron approximation greatly simplifies many-body interactions, the one-electron Schrödinger equation is still complicated to solve because the effective potential depends on all other electron wavefunctions, which are also determined by coupled Schrödinger equations. Thus, the mean-field theory is an obvious choice to overcome this problem. For example, the state-of-the-art method, density functional theory (DFT), is developed to study the electronic structure of many-body systems.

The origin of DFT can be traced back to 1920s, when Thomas and Fermi [3] proposed a total energy functional theory to describe the inhomogeneous electron gas. Instead of the coordinates of electrons, the density of electrons was taken as the variable, which extremely simplified the equation, reducing it from $3N$ spatial
coordinates to 3 spatial coordinates. However, their final result was not accurate enough because of the absence of the exchange and correlation between electrons; the electron exchange arises from anti-symmetry requirements of wavefunctions by the Pauli exclusion principle; the electron correlation consists of the Fermi correlation and Coulomb correlation, while the former prevents two parallel-spin electrons from being found at the same point in space, the latter describes the correlation between the spatial position of electrons due to their Coulomb repulsion.

Despite its inaccuracy, this density-based functional suggested a new direction to solve many-body problem. In 1964, Hohenberg and Kohn published a paper proving that the desired density functional existed [4]. Their key point is that there is a universal functional of the electron density for total energy of a many-electron system in the presence of any given external potential; the ground-state properties are uniquely determined by the electron density, and the correct ground-state electron density minimizes this energy functional. Later, this idea was also expanded to the spin-density functional and finite temperature cases [5,6].

The Hohenberg-Kohn theorem only proved that there existed a density functional, but did not give the details of the form of that functional. In 1965, Kohn and Sham proposed a self-consistent procedure to apply DFT to atomic and solid-state calculations [7]. They included all of the many-body effects into the exchange-correlation energy functional and minimized the energy functional to derive the Kohn-Sham single-particle equations. The many-body effects created a new local exchange-correlation potential term $V_{xc}$ in addition to the Hartree term.
The key problem in DFT is to find a working or approximated form of the exchange-correlation term since no one can give a simple exact expression for it. To date there are two popular approximations, the local density approximation (LDA) and the generalized gradient approximation (GGA). Generally the exchange-correlation energy is an integral over all space of the exchange-correlation energy density weighted by the electron density. The LDA assumes that this exchange-correlation energy density at each point is the same as that of the homogeneous electron gas with the same density. The GGA form proposes a more accurate way to include the gradient term of the local density. A few forms of the expansion of GGA have been proposed, such as Becke (B88) [8], Perdew and Wang (PW91) [9], and Perdew, Burke and Enzerhof (PBE) [10]. With these forms of the exchange-correlation term, the Kohn-Sham equation including exchange-correlation terms can be solved in a self-consistent scheme [7].

Although the self-consistent scheme of the Kohn-Sham equation provides a convenient way to calculate the electronic structure of solids, the direct calculation of the total energy has to deal with divergent terms due to interactions between an infinite number of ions in the crystal. The efficiency of the calculation is also reduced by the computation of a large number of six-dimensional integrals for the e-e interaction terms, making it extremely time consuming and posing a practical difficulty to achieve a desired accuracy. To overcome these difficulties, the momentum-space formalism for total-energy calculations [11] is proposed for crystals. In momentum-space representations, the integral of the Coulomb interaction becomes a sum of rapid convergent
series, and the divergent terms of energies from $e-e$, electron-ion and ion-ion interac-
tions turn out to cancel one another. Upon performing the band-structure calculation,
the total energy is automatically obtained as a sum of a few previously calculated
terms without multicentre integrations, thus greatly simplifying the computation and
achieving the convergence rapidly.

The momentum-space formalism is particularly efficient for combining the pseu-
dopotential method [11] with a plane-wave basis set [12], which can further reduce
computational efforts for calculating electronic structures of solids. The idea of the
pseudopotential approximation is to replace the true atomic potential by a pseudopo-
tential that effectively reproduces the effects of the core electrons, for a reason that
the core-electron wavefunctions remain essentially unchanged when placed into dif-
ferent chemical environments and the core wavefunctions’ only major contribution to
chemical bonding is to enforce the valence wavefunctions’ orthogonality to the core
states. Pseudopotentials can be generated by many different schemes that may result
in different transferability of the pseudopotential and different convergence speeds in
simulations. In 1991, Troullier and Martins proposed a procedure to generate efficien-
t first-principles norm-conserving pseudopotentials, which are exemplified for a wide
range of materials including first-row, transition metal or rare-earth elements [12].
Nowadays, this pseudopotential approximation is routinely used in first-principles
calculations

Despite its simple treatment of exchange-correlation, DFT has been proven to be
a great success in describing the ground-state properties of atoms, molecules, and
solids [10,13,14]. In particular, structural properties of solids such as the lattice constant, the bulk modulus, and the cohesive energy calculated by DFT are generally within a few percent of the experimental values [15,16]. Maybe more importantly, this approach can be easily implemented into the large-scale parallel simulating platforms. As a result, DFT has been a standard practice for first-principles calculations nowadays.

Despite its broad utility, DFT is an intrinsically ground-state theory and it cannot provide reliable descriptions of excited-state properties. For example, the one-particle eigenvalues in DFT have no formal ties to quasiparticle (QP) energies, an excited-state feature, which are necessary for taking account of the response of an electron gas when adding an electron to the system or subtracting one from it. For simple metals, DFT usually gives reasonable overall results for electronic structures since these electrons behave more like free-electron gas, although the zone-edge gaps or bandwidth of some particular metal such as Al and Na do not agree with experiment well [17,18]. For semiconductors and insulators, the discrepancies are much more dramatic. The DFT-calculated band gap (Kohn-Sham gap) is typically 30-50 % less than that observed in the optical spectrum [16]. The discrepancies for the details of the band dispersions are typically smaller, but are also material dependent. For example, the bandwidth of Ge is well represented by the LDA eigenvalues [19,20], but for diamond, it is substantially underestimated as compared to x-ray spectra [21].

The aforementioned shortcomings of DFT predictions are due to deficiencies of the currently used exchange-correlation functionals, which are inadequate for describing
excited-state properties, such as QP energies and optical properties. There are two main reasons for inaccuracies in DFT.

First, the approximated functionals of DFT are incorrectly convex between the integer numbers of charges, violating the linearity of the energy as a function of fractional charges because of the discrete nature of electrons. This results in the so-called delocalization error [22–24]. For the exact functional, its straight-line behavior means that the derivatives of the energy with respect to fractional charge yields the ionization energy and the electron affinity. When adding or removing an electron from a system, approximated functionals produce an overly disperse distribution for the added electron or hole and produce smaller derivatives due to the incorrect convex behavior. This explains the well-known underestimation of the band gap of DFT.

Secondly, the approximated functionals do not accurately describe the interaction between the degenerated spin-up and spin-down states. This introduces a massive error for these fractional-spin states, which is called the static correlation error [24]. This problem arises because of the difficulty in using the electron density to describe the interaction of degenerate states; it can be understood as from the violation of constancy of the energy as a function of fractional spins. This static correlation error is particularly important in strongly correlated systems, as exemplified by Mott insulators, including transition metal oxides, where DFT is unable to predict the correct insulating state.

The above deficiencies of approximate functionals make DFT suffer from systematic problems in describing charge-transfer processes, excitation energies in molecules,
response properties in solids, electron transport, and the band gaps of semiconductors. Therefore, a more accurate description of exchange-correlation interactions is desired. To date several methods with better treatments of dynamical correlations have been developed [25–29]. One of the successful methods is the many-body perturbation theory (MBPT) [29,30], which is based on a set of Green’s-function equations, starting from a one-electron propagator and further considering the electron-hole (e-h) Green’s function for the optical response. The key ingredients of the MBPT are the descriptions of the electron’s self-energy $\Sigma$ and the e-h interaction.

The electron self-energy represents the contribution to the electron energy from interactions between this electron and all other electrons. It contains the effect of exchange and correlation simultaneously. The self-energy operator $\Sigma$ is in general a nonlocal, energy-dependent, non-Hermitian operator. A good approximation for $\Sigma$ is the Hedin’s GW approach [31, 32] by expanding the electron self-energy in a perturbation series in terms of the dynamically screened Coulomb interaction (W) and the dressed Green’s function (G). For this approach, the crucial input is the dielectric matrix which describes the dynamical screening in solids. In practical, the static dielectric matrices are first obtained from the irreducible palorizability by the first-principles DFT approach within the random-phase approximation (RPA) [29]. Then the dielectric matrices can be extended to the finite-frequency dependence with a generalized plasmon-pole (GPP) model [29] or the full frequency calculation [33,34].

Usually, the dielectric function and the Green’s function are constructed by the DFT-calculated wavefunctions and eigenvalues. In particular, experimentally mea-
sured electronic charge distribution is very similar to that from DFT, hinting that DFT calculated wavefunctions are similar to those of QPs [27, 29]. Therefore, this single-shot $G_0W_0$ calculation is good enough to obtain accurate band gaps of usual semiconductors without further self consistent calculations. However, for correlated systems, in which the DFT-calculated wavefunctions and eigenvalues are far away from the realistic cases, the self-consistent GW scheme will be necessary.

The GW approximation on QP self-energy is a powerful tool to determinate QP energies of semiconductors and insulators using DFT as the zero-order starting point for this perturbative calculations. The GW calculations generally yield an excellent agreement between the experiment and the calculated band structure [29, 35–38], apart from the case of strong correlated systems. For example, Hybertsen and Louie calculated diamond, Si and Ge materials with the GW approach, giving the indirect band gap of 5.5, 1.29, 0.75 eV, as compared with experimental gaps of 5.48, 1.17, and 0.744 eV, respectively [29].

In the past thirty years, many studies have been performed to improve our understanding of the first-principles GW methods. Hybertsen and Louie systematically reviewed the GW approximation and discussed the role of local fields in the self-energy operator [29]. Bechstedt discussed the physics of the GW method in relation to the calculation of the dielectric response and the self-energy [39]. Godby reviewed QP calculations for simple metals and semiconductors [40]. Mahan examined different GW approximations by including the self-energy and vertex diagrams beyond the random phase approximation [41]. Pollmann, et al. concentrated on GW calculations
of semiconductor surfaces [42], and Farid considered QP calculations from mathematical aspects [43]. Onida et al. reviewed the prominent approaches for electronic excitations calculations, and compared the GW method with the time-dependent density-functional theory (TDDFT) [27].

The GW approach yields very satisfactory results concerning the band structure, but it is not enough when one is interested in optical spectroscopies, such as the optical absorption, where e-h excitations are essential. In particular, optical spectra are so useful to investigate condensed-matter systems; absorption, reflectivity, photoluminescence, and other optical techniques are commonly used to characterize materials. Additionally optical excitations provide the basis for a vast range of technical applications, including light-emitting devices, laser technology, and photovoltaics. In this context, it is of great importance to describe accurately such excitations by highly reliable and efficient \textit{ab initio} approaches.

The investigation of optical e-h excitation requires an effective two-body approach that goes beyond the single-particle picture of individual electron and hole excitations and includes a proper treatment of the e-h interaction. The key point is to investigate the two-particle Green’s function and solve its equation of motion, which is know as the Bethe-Salpeter equation (BSE) [30]. The two-body Green’s function $G_2$ is constructed on the basis of the one-body Green’s function $G_1$, which may be described by the GW approximation. Starting from the QP electron and hole states of $G_1$ and their QP energies, the e-h interaction, which results from the self-energy operator, is calculated. Thereafter, the BSE equation is solved, yielding the coupled,
correlated $e$-$h$ excitation states. Together with the corresponding optical transition matrix elements that result as coherent superpositions from those of free $e$-$h$ pairs, the entire linear optical spectrum of a material can be obtained.

The first-principles MBPT (GW+BSE) used currently mostly follows the scheme first introduced by Onida for addressing the spectrum of a $Na$ cluster and that of the insulator $Li_2O$ [44]. This procedure consists of three major steps: (i) a ground-state DFT calculation; (ii) a GW calculation to obtain the dynamic screening and corrected QP energies; (iii) solving the $e$-$h$ interaction in BSE and calculating the optical spectra. This scheme has been applied to calculate the absorption spectra of many materials, including molecules [30], small clusters [30, 44], polymers [45], and periodic crystals of semiconductors and insulators [30, 46, 47]. These simulation results agree with the experimental data quite well [30,44–47].

In addition to the electronic structures and electron excitations, the lattice vibrations, phonons, are also very interesting and important, since they are the foundation for understanding thermal properties of solids. With the help of the electronic structure calculations that describe well the ion-electron interactions, phonons, a QP concept of an excited state of the collective atomic vibrations, can be accurately calculated by DFT-based techniques studying the response of crystals to external perturbations.

There are mainly two types of phonon calculation methods based on DFT. The first one is the small-displacement method which is also referred as the “direct” method [48, 49]. The calculation is based on the fact that the force constant ma-
trix (or the Hessian matrix) expresses the proportionality between forces and displacements, and the relation is linear when the displacements are small enough. The method, in principle, is to displace a single atom in a cell along a Cartesian direction, with all other atoms being fixed at their equilibrium positions; the resulting forces on all the atoms give directly the elements of the force constant matrix. If this procedure is repeated for every single atom along every direction, all the elements of the force-constant matrix can be obtained, and then the phonon frequencies can be calculated by diagonalizing the matrix. The small-displacement method is rather straightforward computationally, and it is easy to be implemented in the total-energy code. Additionally, this method includes both linear and nonlinear effects directly. However, the convergence can be slow in periodic ionic materials, and the method is computationally expensive for a system containing a large number of atoms per unit cell.

The other DFT-based phonon calculation method is the linear-response method [50]. The main idea is from the Hellmann-Feynman theorem that a linear order variation of the electron density upon a perturbation applied to the crystal is responsible for a variation of the total energy up to the second order of the perturbation. If the system is perturbed in a phonon wave with a wave-vector \( q \), the corresponding change in electron density to the first order can be obtained using only unperturbed wavefunctions by the perturbation theory, and it can further be used to determine the dynamical matrix (DM) at this wave vector \( q \). Therefore, only the ground-state calculations of crystal are required, and this procedure can be done for any arbitrary
wave vector, without the need for the construction of a supercell. As in its name, the linear-response method only considers the linear effect of the perturbation, and it needs extra steps to calculate higher-order terms for the nonlinear effect. The merits of this method lie particularly in the phonon calculations of semiconductors and insulators, in which the forces induced by perturbations are relative long-range. The linear-response method does not need to construct a supercell and overcomes the slow convergence problem in the limit of zero wave-vector, which significantly increases the efficiency of the computation.

The bottleneck of the linear-response method involves a large number of electronic ground states for calculating the linear-order change of electron density. Therefore, several new developments have been proposed to advance the linear response method [51–54]. One of the developments uses a Green’s function method [52, 53, 77] and avoids the requirement of inclusion many extra conduction states, greatly improving the computation efficiency. Since the 1980’s, first-principles linear-response methods have been successfully used to study phonons in many materials. They obtain satisfying results about lattice vibrations and have been even extended to study the nonlinear effect.

From the brief review above, there is no doubt that first-principles calculations have developed to become powerful tools for studying the electronic, optical, and thermal properties of a wide range of materials. For the rest of this chapter, we will introduce how we apply these state-of-the-art approaches in studying electronic
structures, lattice vibrational modes, and optical excitations of novel materials as motivated by recent experiments.

In the past twenty years, reduced-dimensional materials have attracted tremendous interest not only because new physics and exotic properties emerge in these materials due to quantum confinement effects and the large surface-volume ratio, but also because the manipulation and engineering of matters on an atomic scale or nanometer size yield promising applications in various fields including nanoelectronics, nanobiotechnology, environments, and renewable energy. Numerous types of reduced-dimensional materials have been fabricated; examples of them include nanoclusters, nanowires, nanotubes, nanoribbons, graphene and graphene derivatives. In these materials, the size along at least one dimension is reduced to the nanometer scale, much smaller than other remaining dimensions. At this point, the electronic interactions and phonon vibrations of these reduced-dimensional materials are significantly changed by quantum confinement, making it hard to apply empirical models based on bulk materials to them. In this sense, first-principles calculations are valuable to study the properties of reduced-dimensional materials because of their parameter-free character.

In Chapter 2, I will first provide a detailed discussion about the first-principles methods used in the studies of reduced-dimensional materials. The first is about DFT as the foundation of electronic-structure and phonon calculations. The second is the MBPT including the GW approximation and BSE for calculating QP energies,
excitons, and optical absorptions. The third one is the linear-response theory for the first-principles calculations of the lattice vibrational modes and phonons.

The first research topic of my PhD work is the band offsets engineering in Si/Ge core-shell nanowires (NWs) presented in Chapter 3. Heterojunctions are of fundamental interest in condensed matter physics and are also elementary building blocks of microelectronics. Recently nanoscale radial heterojunctions in one-dimensional (1D) core-shell NWs have been fabricated and have shown promising electrical and optical properties associated with broad applications. Compared to traditional planar bulk heterojunctions, core-shell NWs have more degrees of freedom to tune the band offsets, such as variations in the geometry and chemical components of core and shell. Therefore, understanding the electronic structure and, more importantly, how to control band alignments are of central interest for researchers of core-shell nanostructures. In this chapter, we will apply DFT to show that the applied axial strain and the variation of core-shell size can modify the band offset of an important type of core-shell nanostructures, the Si/Ge core-shell NWs.

In the second project, we present a first-principles study on lattice vibrational modes of Si/Ge core-shell NWs. Recent research has shown that Si/Ge core-shell NWs possess a number of unusual thermal transport features for energy applications. In order to realize and optimize the above interesting properties, it is imperative to determine how to obtain structural information and, thereafter, how to design core and shell geometries. However, compared with bulk semiconductors, there have been very limited first-principles studies on vibrational modes of core-shell NWs. How
quantum confinement and internal strain in core-shell NWs affect their phonons and thermal properties has not been well understood yet. Therefore, calculating lattice vibrational modes in Si/Ge core-shell NWs will be an appropriate starting point to understand the electric and thermal properties of core-shell NWs. In this project, we reveal that the lattice vibrational modes and their frequencies are strongly affected by the geometry of core-shell NWs. The origin of these variations can be attributed to the internal strain due to the lattice mismatch.

In the third project, we study the exciton spectra of two-dimensional (2D) semiconducting carbon structures. The exciton spectrum, the sequence of $e-h$ binding energies, is the most direct way to understand excitonic effects of semiconductors. It is also the foundation for constructing useful models widely used to identify excitonic effects in optical spectroscopy experiments. For example, the $e-h$ binding energy can be conveniently extrapolated from the measured sequence of exciton peaks according to model predictions. In particular, $e-h$ interactions are known to be dramatically enhanced in reduced dimensional structures. How these unique quantum confinements influence exciton spectra and how one subsequently modifies corresponding models have been of fundamental interest.

Recently, many-electron effects and optical properties of graphene and its derivatives have garnered substantial interest because of their unique many-electron effects. However, other than studies of the optical absorption, the exciton spectra of these novel materials are largely unknown. This motivates us to apply the first-principles simulation based on the MBPT to study the exciton spectra of the 2D semiconduct-
ing carbon structures and build an appropriate exciton models that can be applied to generally confined 2D systems.

In the last part of my PhD work, we study the electronic structures and QP energy of silicene structures, a layer of graphene-like hexagonal lattices, which have attracted tremendous interest because of their expected unique linear energy-momentum dispersion similar to graphene. Unlike graphene, a unique character of silicene is that the simply buckled silicene structure is predicted to be more stable than the perfectly planar one.

To date several new silicene structures with more complicated buckling styles are proposed by recent experiments. However, other than interpreting atomistic structures, very limited attempts have been carried out to discover their electronic structures, which are crucial for guiding experimental measurements and understanding the electric and optical properties of these materials of ever-growing interest. Moreover, the buckling of silicon atoms may break the symmetry of the honeycomb lattice. According to the present understanding of graphene, the variation of structure and symmetry shall not only change the Fermi velocity but also possibly generate a finite band gap, promising a crucial advantage for silicene over graphene for broader applications. In this study, we will employ both DFT and GW approximation to calculate the band structure of these novel silicenes. On one hand, our calculation confirms the Dirac-cone electronic structure of planar and simply buckled silicenes; on the other hand, our calculation shows that those newly proposed silicenes are actually
semiconductors instead of graphene-like semimetals. Our further calculations show the different band structures are due to the symmetry breaking.
2. THEORETICAL METHODS

In this chapter, I will give a brief review of the first-principles methods that I use in my study of reduced dimensional semiconductors. This chapter is organized as follows. The first section is an introduction of density functional theory (DFT), which is the foundation of electronic-structure calculations. The proof of the Hohenberg-Kohn theorem is provided and the Kohn-Sham equations are derived under the local density functional approximation (LDA). The pseudopotential method and the momentum-space formalism are discussed as well because they are essential ingredients of modern first-principles simulations based on DFT. The second section is about many-body perturbation theory (MBPT) including the GW approximation for quasiparticle (QP) energies, and the Bethe-Salpeter equation (BSE) for excitons and corresponding optical absorption spectra. The single-particle Green’s function method including dynamic screening is employed to calculate the electron self-energy and thereby correct the QP energies. By using the two-body Green’s function, the BSE containing electron-hole (e-h) interactions are solved to obtain excitonic states. The optical absorption spectrum is then evaluated by a coherent summation of transition-matrix elements of those contributing e-h pairs. The last section is about the first-principles calculation of phonons and we specially focus on the linear response theory, in which the dynam-
2.1 Density Functional Theory (DFT)

DFT is a mean-field treatment of many-body systems. It is based on the Hohenberg-Kohn theorem that the electron density can uniquely determine the ground-state properties of a many-electron system in the presence of any given external potential. The exchange and correlation of many-electron interactions are included in an approximated functional, and the solution of the many-electron problem is then reduced to a self-consistent solution of a set of effective one-particle equations (Kohn-Sham equations). This method has the great advantage for being a practical computational tool to handle solids when combined with the pseudopotential method under the momentum-space representation.

2.1.1 Hohenberg-Kohn Theorem

The Hohenberg-Kohn theorem [4] states that there is a universal functional of the electron density, $E[n(\vec{r})]$, for the total energy of a many-electron system in the presence of any given external potential $v(\vec{r})$; the ground-state properties are uniquely determined by the electron density $n(\vec{r})$, and the correct ground-state electron density $n_0(\vec{r})$ minimizes this energy functional $E[n(\vec{r})]$. 
Following Hohenberg and Kohn’s work, this theorem can be proven in two steps [4]. First, we have to show that the potential $v(\vec{r})$ is a unique functional of the electron density $n(\vec{r})$.

The Hamiltonian of the system has the general form of

$$H = K + V + U,$$  \hspace{1cm} (2.1)

where $K$ is the kinetic energy, $U$ is the interaction energy between electrons, and $V$ is the external potential energy. For simplicity, we are only dealing with the situations, in which the ground state is non-degenerate. Since the Hamiltonian $H$ is fixed for a given external field $v(\vec{r})$, the charge density $n_0(\vec{r})$ of the ground state $|\Psi\rangle$ is clearly a unique functional of $v(\vec{r})$.

We then need to show, conversely, that the $v(\vec{r})$ is a unique functional of $n_0(\vec{r})$, apart from a trivial additive constant. The proof is proceeded by argumentum ad absurdum.

Assume another, different, potential $v'(\vec{r})$ with the ground state $|\Psi'\rangle$, gives rise to the same density $n_0(\vec{r})$. Because of the minimal property of the ground state, we get

$$E' = \langle \Psi' | H' | \Psi' \rangle < \langle \Psi | H' | \Psi \rangle = \langle \Psi | H + V - V | \Psi \rangle,$$  \hspace{1cm} (2.2)

that is

$$E' < E + \int [v'(\vec{r}) - v(\vec{r})] n_0(\vec{r}) d\vec{r}.$$  \hspace{1cm} (2.3)
By interchanging primed and unprimed quantities, we find

\[ E < E' + \int [v(\vec{r}) - v'(\vec{r})] n_0(\vec{r}) d\vec{r}. \]  

(2.4)

Addition of Eq. (2.3) and Eq. (2.4) leads to the inconsistency

\[ E + E' < E' + E. \]  

(2.5)

Therefore, the external potential \( v(\vec{r}) \) is a unique functional of the ground-state charge density \( n_0(\vec{r}) \).

Secondly, there is a universal functional \( F[n_0] \) for describing the kinetic and interaction energy, so that the ground-state charge density \( n_0(\vec{r}) \) uniquely determines the minimization of the total-energy functional of the many-electron system.

Since \( v(\vec{r}) \) is a unique functional of \( n_0(\vec{r}) \), the ground state \( |\Psi\rangle \) is a functional of \( n_0(\vec{r}) \), if we define

\[ F[n_0(\vec{r})] \equiv \langle \Psi | K + U | \Psi \rangle. \]  

(2.6)

Then \( F[n_0] \) is a universal functional of \( n_0(\vec{r}) \), valid for any number of particles and any external potential.

For a given potential \( v(\vec{r}) \), the total-energy functional is define as

\[ E[n_0] \equiv \int v(\vec{r}) n_0(\vec{r}) d\vec{r} + F[n_0]. \]  

(2.7)
Clearly, $E[n_0]$ is a universal functional of $n_0(\vec{r})$; for the correct $n_0(\vec{r})$, $E[n_0]$ equals the ground-state energy.

Under the charge conservation condition, $\int n_0(\vec{r})d\vec{r} = N$, the total-energy functional $E[n_0]$ should have its minimum value only for the correct density $n_0(\vec{r})$. Consider any other charge density $n'_0(\vec{r})$, we have

$$E[n'_0] = \int v(\vec{r})n'_0(\vec{r})d\vec{r} + F[n'_0] > E[n_0] = \int v(\vec{r})n_0(\vec{r})d\vec{r} + F[n_0].$$  \hspace{1cm} (2.8)$$

Thus the minimization of the total-energy functional gives the correct ground-state charge density.

The Hohenberg-Kohn theorem has been proven. Moreover, if $F[n_0]$ were a known and sufficiently simple functional of $n_0(\vec{r})$, the ground-state problem of the many-electron system in a given external potential can be easily solved since it requires merely the minimization of a total-energy functional according to the charge density. However, the major complexities of many-electron problems are associated with the determination of the universal functional $F[n_0]$.

### 2.1.2 Kohn-Sham Equations

Based on the Hohenberg-Kohn theorem, Kohn and Sham developed DFT further [7]. They used the single-particle approximation and introduced an effective potential including the exchange and correlation effects. The equation of many-electron systems transforms to a set of one-particle Schrödinger equations, which can be solved
efficiently in the self-consistent scheme. In the following part of this subsection, we will show the major steps of the Kohn-Sham method.

From the Hohenberg-Kohn theorem, the ground-state energy of an interacting electron gas in a static external potential \( v(\vec{r}) \) can be written in the form

\[
E = \int v(\vec{r})n(\vec{r})d\vec{r} + \frac{1}{2} \int \int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|}d\vec{r}d\vec{r}' + G[n],
\]

(2.9)

where \( n(\vec{r}) \) is the charge density and \( G[n] \) is a universal function of the density.

The first approximation for \( G[n] \) is to split out the kinetic energy part by writing

\[
G[n] \equiv K[n] + E_{xc}[n],
\]

(2.10)

where \( K[n] \) is the kinetic energy of a system of noninteracting electrons with the density \( n(\vec{r}) \), and \( E_{xc}[n] \) is the exchange-correlation energy of an interacting system with the same density \( n(\vec{r}) \). The term \( E_{xc}[n] \) contains all the many-body effects except for the Hartree term, although the exact form of \( E_{xc}[n] \) is unknown. However, an approximation of the exchange-correlation energy can be applied for a slowly varying \( n(\vec{r}) \) as below:

\[
E_{xc}[n] = \int n(\vec{r})\epsilon_{xc}(n(\vec{r}))d\vec{r},
\]

(2.11)

where \( \epsilon_{xc}(n(\vec{r})) \) is the exchange-correlation energy per electron of a uniform electron gas of density \( n \). This is called the local density approximations (LDA), in which
the exchange-correlation energy is determined by the local term of the charge density functional.

Subject to the condition of a fixed number of electrons, one can apply the variational principle to Eq. (2.9) and gets:

\[ \int \delta n(\mathbf{r}) \left\{ \varphi(\mathbf{r}) + \frac{\delta K[n]}{\delta n(\mathbf{r})} + \mu_{xc}(n(\mathbf{r})) \right\} d\mathbf{r} = 0; \]  

(2.12)

here the effective potential has two parts:

\[ \varphi(\mathbf{r}) = v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r'}|} d\mathbf{r'}, \]  

(2.13)

and

\[ \mu_{xc}(n) = \frac{d(n\epsilon_{xc}(n))}{dn}. \]  

(2.14)

\( \mu_{xc}(n) \) is the exchange and correlation contribution to the chemical potential of a uniform gas of the density \( n \).

Then Eq. (2.12) results in a set of single-particle Schrödinger equations (in Hartree atomic units)

\[ \left\{ -\frac{1}{2} \nabla^2 + [\varphi(\mathbf{r}) + \mu_{xc}(n(\mathbf{r}))] \right\} \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}), \]  

(2.15)

and the density is

\[ n(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2, \]  

(2.16)

where \( N \) is the number of electrons.
Eq. (2.15) is also called the Kohn-Sham equation. It significantly simplifies the many-body problem by including exchange and correlation as approximated functionals and reducing to independent-particle equations with new effective potentials \( \varphi(\vec{r}) + \mu_{xc}(n(\vec{r})) \).

The equations from Eq. (2.13) to Eq. (2.16) provide a self-consistent procedure for practical calculations, as shown in Figure 2.1: one begins with a guess of the initial density \( n(\vec{r}) \), constructs \( \varphi(\vec{r}) \) from Eq. (2.13) and \( \mu_{xc}(n(\vec{r})) \) from Eq. (2.14), solves Eq. (2.15) to get eigenvalues and eigenfunctions, then constructs a new density by Eq. (2.16), updates the equations, and repeats the procedure until the solutions converge. After the convergence, the energy of the system is given by

\[
E = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \int \int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' + \int n(\vec{r})[\varepsilon_{xc}(n(\vec{r})) - \mu_{xc}(n(\vec{r}))] d\vec{r}, \tag{2.17}
\]

where \( \varepsilon_i \) is the eigenvalue of Kohn-Sham particles in Eq. (2.15).

The most essential part the Kohn-Sham equations is to determine the approximated functional of exchange and correlation effects between electrons, which decides the accuracy in describing the many-electron interactions. In order to get an explicit form of the approximated functional, the exchange term can be analytically calculated with the Hartree-Fork approximation, while the correlation part is obtained with some other methods, such as the quantum Monte Carlo method (QMC), for LDA. Several models, including the “Ceperley-Alder” form (Pedew-Zunger (PZ)) [55, 56], are provided for fitting the functional form from the simulation results. Besides LDA,
Figure 2.1. Schematic representation of the self-consistent loop for solving Kohn-Sham equations.
another widely used approximated functional is the generalized gradient approximation (GGA), which includes the gradient term of the local density and could improve the approximation for some materials with rapidly varying charge density [8–10]. In practical calculations, one can choose the type of functionals depending on the details of the system studied.

DFT can only describe well the ground state of electronic systems because the solution of DFT is based on variational principles. Meanwhile, the exchange-correlation approximated functionals suffer deficiencies in describing many-body interactions between electrons [22–24]. Therefore, the eigenvalues and eigenfunctions of the Kohn-Sham equation, Eq. (2.15), can not be directly understood as QP states of a many-body system. They actually correspond to the states of Kohn-Sham particles, a collection of fictitious particles. Although the sum of the modula of Kohn-Sham eigenfunctions gives very good results in comparison with the physical density of the system, the Kohn-Sham eigenenergies fail in describing the QP energies accurately. For example, DFT with LDA or GGA usually underestimates the QP band gap of semiconductors [16]. However, this method is widely used in a lot of calculations, because to people’s surprise, in most cases, the band structure of the Kohn-Sham particles can at least qualitatively explain experimental results, and in some case, it even gives quantitatively good results compared to experiments. As a result, despite of the deficiencies of Kohn-Sham method, it has been a standard way to calculate the electronic band structure with the Kohn-Sham equation as the first estimation.
2.1.3 Pseudopotential Method

Although DFT greatly simplifies the complexities of electronic-structure calculations of solids, the magnitude of the computation effort is still enormous, specially for complicated crystals with many-core-electron atoms. The pseudopotential method is particularly useful to further reduce the number of the electrons involved. This idea is to replace the true atomic potential by a pseudopotential that effectively reproduces the effects of the core electrons; then we only need to solve equations of valence electrons. The pseudopotential approximation is justified because the core-electron wavefunctions remain essentially unchanged when placed into different chemical environments and the core wavefunctions’ only major contribution to chemical bondings is to enforce the valence wavefunctions being orthogonal to the core states [12].

Another advantage of the pseudopotential method is related to the plane-wave-basis formalism employed. The plane-wave basis is a simple and stable way to expand electron wavefunctions in crystals [11]. However, expanding the core wavefunctions or the core oscillatory region of wavefunctions into plane waves is extremely inefficient because of their deep and sharply-varying features [12]. The problem can be solved by combining the plane-wave-basis formalism with pseudopotentials. The pseudopotential and plane-wave method is highly efficient in simulations because the number of plane waves needed to expand the electronic states is greatly reduced.

The early work of pseudopotentials can be found in the references [57,58]. Pseudopotentials are mainly generated by two ways. One is to use empirical parameters to
fit known experimental results and construct fitted pseudopotentials for new-property
calculations; this is called the empirical pseudopotential [59,60]. The other way is to
calculate all-electron wavefunctions of an atom through the first-principle approach
and construct the pseudopotential without fitting experiments, this type is called the
\textit{ab initio} pseudopotential. The \textit{ab initio} pseudopotentials have good transferability
and accuracy, and have been widely used in first-principles calculations. Depending
on how to construct pseudopotential from all-electron calculation, \textit{ab initio} pseudo-
dopotentials can be generated by different schemes. In my work, most \textit{ab initio}
pseudopotentials are of the Troullier-Martin form [12].

The Troullier-Martin pseudopotential belongs to the norm-conserving pseudopo-
tential, a type of \textit{ab initio} pseudopotentials. Norm-conserving pseudopotentials en-
force the condition that, outside of a cutoff radius, the norm of each pseudo-wavefunction
is identical to its corresponding all-electron wavefunctions [61].

For all-electron atomic calculations, the angular-momentum dependent spherical-
form of the Kohn-Sham equation can be written as

\[
\left[\frac{-1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V(\rho; r)\right]r R_m(r) = \varepsilon_m r R_m(r). \tag{2.18}
\]

\(V(\rho; r)\) is the self-consistent one-electron effective potential.

\[
V(\rho; r) = \frac{-Z}{r} + V_H(\rho; r) + V_{\text{XC}}^{\text{LDA}}(\rho(r)), \tag{2.19}
\]
where $\rho(r)$ is the sum of the electron density for the occupied wavefunction $R_{nl}(r)$, $V_H(\rho; r)$ is the Hartree potential, and $V_{XC}^{LDA}(\rho(r))$ is LDA for the exchange-correlation potential.

Norm-conserving pseudopotentials are constructed such that they satisfy the following four general conditions [12]:

1. The valence pseudo-wavefunctions generated from the pseudopotential contain no nodes. This is because we would like to obtain smooth pseudo-wavefunctions which can be expanded by a small number of plane waves.

2. The normalized atomic radial pseudo-wavefunction (PP) is equal to the normalized radial all-electron wavefunction (AE) beyond a chosen cutoff radius $r_{cl}$,

$$R_{i}^{PP}(r) = R_{i}^{AE}(r), \quad \text{for} \quad r > r_{cl},$$

or converges rapidly to that value.

3. The charge enclosed within $r_{cl}$ for the above two wavefunctions must be equal,

$$\int_{0}^{r_{cl}} |R_{i}^{PP}(r)|^2 r^2 dr = \int_{0}^{r_{cl}} |R_{i}^{AE}(r)|^2 r^2 dr.$$  \hspace{1cm} (2.21)

4. The valence all-electron and pseudopotential eigenvalues must be equal,

$$\varepsilon_{i}^{PP} = \varepsilon_{i}^{AE}.$$  \hspace{1cm} (2.22)
Once the pseudo-wavefunctions are obtained, the screened pseudopotential is then recovered by the inversion of the radial Kohn-Sham equation (Eq. (2.18)),

$$V_{\text{scr},l}^{PP}(r) = \varepsilon_l - \frac{l(l+1)}{2r^2} + \frac{1}{2r R_l^{PP}(r)} \frac{d^2}{dr^2} [r R_l^{PP}(r)].$$  \hspace{1cm} (2.23)

In order to get the pseudopotential contributed from the nucleus and core electrons, the ion pseudopotential is calculated by subtracting the Hartree $V_H^{PP}(r)$ and exchange-correlation $V_{xc}^{PP}(r)$ potential from the valence pseudo-wavefunctions with the screened pseudopotential,

$$V_{\text{ion},l}^{PP}(r) = V_{\text{sc},l}^{PP}(r) - V_H^{PP}(r) - V_{xc}^{PP}(r).$$  \hspace{1cm} (2.24)

Thus the pseudopotential for the valence-electron calculation is obtained.

As one can see from Eq. (2.23), the pseudopotentials $V_{\text{ion},l}^{PP}(r)$ are different for different angular-momenta $l$. This angular-momentum dependence makes the pseudopotential nonlocal, or more precisely semilocal. Thus the ionic pseudopotential operator can be separated into the local and nonlocal parts:

$$V_{\text{ion}}^{PP}(r) = V_{\text{ion,local}}^{PP}(r) + \sum_l V_{\text{nonlocal},l}(r) \hat{P}_l.$$  \hspace{1cm} (2.25)
The key part of the norm-conserving pseudopotentials generation is the construction of the radial pseudo-wavefunction (PP). Kerker [63] first defined the pseudo-wavefunction as

\[ R^{PP}_l = R^{AE}_l \quad \text{when} \quad r > r_{cl}, \]  

(2.26)

and

\[ R^{PP}_l = r^l \exp(p(r)) \quad \text{when} \quad r < r_{cl}, \]  

(2.27)

where \( p(r) \) is a polynomial of degree \( n = 4 \),

\[ p(r) = c_0 + \sum_{i=2}^{n} c_i r^i. \]  

(2.28)

Troullier and Martins tested the choices of Eq. (2.28) and found that the best results were obtained when two derivatives at the cutoff radii \( r_{cl} \) matched and all odd coefficients of \( C_n \) were set to be zero [12]. They proposed an improved form for \( p(r) \) as

\[ p(r) = c_0 + c_2 r^2 + c_4 r^4 + c_6 r^6 + c_8 r^8 + c_{10} r^{10} + c_{12} r^{12}. \]  

(2.29)

The most important parameter in the above process of the norm-conserving pseudopotentials generation is the cutoff radius \( r_{cl} \), which determines the accuracy and the transferability of the pseudopotential [12,64]. Since the pseudo-wavefunction matches the all-electron wavefunction outside \( r_{cl} \), a smaller cutoff radius can be used to generate pseudopotential with a higher transferability, which means the pseudopotential reproduces other all-electron calculations in different environments more accurately.
However, a smaller cutoff radius results in harder (less smooth) pseudopotentials, which requires a higher energy cutoff or a larger number of plane waves to expand wavefunctions [12,64]. Therefore, it is important to find a balanced cutoff radius in first-principles calculations.

Figure 2.2. Comparison of ionic pseudopotentials for carbon (dotted line for s and solid line for p) in real space (upper part) and reciprocal space (lower part). From left to right the pseudopotentials are generated by schemes of: Troullier and Martins (TM); Kerker; Hamann, Schluter, and Chiang (HSC); Vanderbilt. [12]

Troullier-Martins pseudopotentials are designed to be smooth and they are extremely efficient in plane-wave calculations. Figure 2.2 shows the comparison of carbon pseudopotentials generated by several different schemes [12]. As the figure illustrates, the Troullier-Martins form is the most smooth and thus requires least
plane-wave cutoff energy to achieve the same accuracy of total energy calculations [64]. Moreover, Troullier-Martins pseudopotentials are known to have good transferability for a wide range of elements [12].

Another type of pseudopotentials that I use in my work is the ultrasoft pseudopotential [62], which is implemented into the Vienna Ab-Initio Simulation Package (VASP). This kind of pseudopotential is not norm-conserved, because it does not require the condition that the charge within the cutoff radius $c_l$ is equal to the all-electron result. The advantage of ultrasoft pseudopotentials is that they are smoother (softer) than many other kinds of pseudopotentials and allow the simulation to run extremely quickly. The disadvantage is that they need an extra procedure to compensate the charge difference, making it hard for more advanced calculations, such as the excited-state calculation.

2.1.4 Momentum-Space Formalism

For a crystal structure with a lattice vector $R_\mu$ and ions with a valence $Z$, the corresponding one-electron Schrödinger equation (in Hartree atomic units, Eq. (2.15)) becomes

$$\left\{ -\frac{1}{2} \nabla^2 + \sum_{\mu,l} U_{ps,l}(r - R_\mu) \hat{P}_l + \int \frac{\rho(r')}{|r - r'|} dr' + \mu_{xc}(r) \right\} \psi_i(r) = \varepsilon_i \psi_i(r),$$

(2.30)

where $\sum_{\mu,l} U_{ps,l}(r - R_\mu) \hat{P}_l$ are angular-momentum-dependent pseudopotentials ($\hat{P}_l$ is the projection operator on angular momentum $l$).
Let the momentum-space representations of the wavefunction, the charge density, the inter-electronic Coulomb potential, and the exchange-correlation potential be denoted by $\psi({\bf k}_i + {\bf G})$, $\rho({\bf G})$, $V_{\text{Coul}}({\bf G})$ and $\mu_{\text{xc}}({\bf G})$, respectively, where $G$ are reciprocal lattice vectors.

The momentum representation of the one-electron Schrödinger equation becomes

$$\sum_{G'}[(k_i + G')^2\delta_{GG'} + V_{GG'}^i]\psi(k_i + G') = \varepsilon_i\psi(k_i + G),$$

(2.31)

where

$$V_{GG'}^i = V_{\text{Coul}}(G' - G) + \mu_{\text{xc}}(G' - G) + S(G' - G)[U_{ps}(G' - G) + \sum_l U'_{ps,l,G,k_i+G,k_i+G'}];$$

(2.32)

$S(G' - G)$ is the structure factor, and the purely local part of pseudopotentials ($U_{ps}(G)$) and the non-local part ($U'_{ps,l,G,k_i+G,k_i+G'}$) are treated separately [11].

After solving Eq. (2.31), the total energy of the crystal is given by

$$E_{\text{total}} = \sum_i \varepsilon_i - \frac{1}{2} \int \frac{\rho({\bf r})\rho'({\bf r'})}{|{\bf r} - {\bf r'}|} d{\bf r} d{\bf r'} - \frac{1}{4} \int \mu_{\text{xc}}({\bf r})\rho({\bf r}) d{\bf r} + \frac{1}{2} \sum_{\mu,\nu,\mu' \neq \nu} \frac{Z^2}{|{\bf R}_\mu - {\bf R}_\nu|}$$

$$= \sum_i \varepsilon_i - \Omega \left[ \frac{1}{2} \sum_{G} V_{\text{Coul}}(G)\rho(G) + \frac{1}{4} \sum_{G} \mu_{\text{xc}}(G)\rho(G) \right] + \frac{1}{2} \sum_{\mu,\nu,\mu' \neq \nu} \frac{Z^2}{|{\bf R}_\mu - {\bf R}_\nu|}$$

(2.33)
where $\Omega$ denotes the total volume of the system, and the last term in Eq. (2.33) is the ion-ion Coulomb energy. We define the volume per atom as $\Omega_{at} = \Omega/N$, in which $N$ is the total number of atoms.

Though the ion-ion Coulomb energy term is divergent, it could be cancelled by the long-range divergent part of the Coulomb energy of electrons (the second term in Eq. (2.33) when $G=0$). Then the final expression for the total energy per atom becomes

$$E_{total\,(per\,atom)} = \frac{1}{N} \sum_{i} \varepsilon_i - \frac{1}{2} \Omega_{at} \sum_{G \neq 0} V_{Coul}(G) \rho(G) - \frac{1}{4} \Omega_{at} \sum_{G} \mu_{xc}(G) \rho(G) + \alpha_1 Z + \gamma_{Ewald}. \quad (2.34)$$

The two coefficients are

$$\alpha_1 \equiv \lim_{G \to 0} \left[ U_{ps}(G) + \frac{4\pi Z^2}{\Omega_{at} G^2} \right] = \frac{1}{\Omega_{at}} \int \left[ U_{ps}(r) + \frac{Z}{r} \right] dr, \quad (2.35)$$

and

$$\gamma_{Ewald} = \frac{1}{2} \sum_{\nu} \frac{Z^2}{|R_{\nu}|} - \frac{1}{2} \lim_{G \to 0} \frac{4\pi Z^2}{\Omega_{at} G^2} = \frac{1}{2} \left( \sum_{\nu} \frac{Z^2}{|R_{\nu}|} - \frac{1}{\Omega_{at}} \int \frac{Z^2}{r} dr \right). \quad (2.36)$$

where the prime means that $R_{\nu} = 0$ is excluded in the summation.

The momentum-space formalism with the pseudopotential scheme greatly enhances the convergence rate of total-energy calculations by replacing the integrals of Coulomb interactions in the real space with summations of series in the reciprocal space. As a result, it is widely implemented in current first-principles calculations.
2.2 Many-Body Perturbation Theory (MBPT)

DFT provides a concise way to incorporate exchange and correlation effects into the calculation of the ground-state energy of the interacting-electron system in an external potential. However, a lot of interesting phenomena such as photoemission, electron energy loss spectroscopy, and optical absorption, etc., are closely related to the excited states of electrons instead of their ground state. DFT is inadequate to describe these excited properties because of its systematic deficiencies. If the electronic system is not strongly correlated, most of the above phenomena are related to the electronic states close to the Fermi level, making it possible to describe the collective behavior of the excited states with the QP concept.

The basic idea of the QP concept is used to describe motions in a strongly interacting many-body system. Motion in a solid is extremely complicated, for example, one electron’s motion will push and pull other electrons in the system via Coulomb interactions, implying a redistribution of charge (wavefunctions). In other words, depletion of charge (holes) or accumulation of charge (electrons) in some places, and the response of other electrons conversely affects the electron’s motion. In order to describe the collective behavior of this system, a QP concept is introduced as a fictitious non-interacting (or weekly interacting) particle which contains the effects of the electron’s motion and other electrons’ responses together.

MBPT has been successfully applied to study QP energies and excited-state properties of many-body systems. MBPT is based on a set of Green’s-function equation-
s, starting from a one-electron propagator and further considering the e-h Green’s function for the optical response. MBPT mainly consists of two parts: the GW approximation and the BSE methodology.

The Green’s function approach provides a rigorous formulation for the QP properties. Within the context of the one-particle Green’s function, QPs are associated with the peaks in the spectral function. The self-energy of QPs, which includes the exchange and correlation among electrons, can be expanded by a perturbation series in terms of the dynamically screened Coulomb interaction (W) and the dressed Green’s function (G). The GW approximation only considers the first order of this expansion, in which the self-energy (Σ) is the product of the one-particle Green’s function (G) and the screened interaction (W), Σ = iGW. Then the self-energy operator is used to substitute the exchange-correlation approximate functionals in the Kohn-Sham equations, and the QP energies are calculated by solving the new QP equations. The GW method has been very successful in determining QP energies of semiconductors and insulators by using DFT as the zero-order starting point for these perturbation calculations.

When dealing with optical absorption of semiconductor materials, the above GW is not enough. When a photon is absorbed by the material, an electron is excited to the conduction band, leaving a vacancy in the valence band and creating a QP called the hole. The excited electron and hole interact with each other and form a kind of charge-neutral excitation, which is called the exciton or e-h pair. In this case, the single-particle Green’s function is not enough to describe the exciton, and
two-particle Green’s function has to be constructed. The equation on the motion of two-particle Green’s function is called the BSE [68–70]. The derivation of BSE can be thought as the second order expansion of the self-energy with the screened interaction. The solutions of the BSE provide eigen-energies and eigen-wavefunctions of exciton states, based on which optical absorption can be obtained by calculating the coherent superpositions of optical transition matrix elements.

In principle, the first-principles MBPT (GW+BSE) can be summarized as the following three steps: (i) a ground-state DFT calculation; (ii) a GW calculation to obtain the dynamic screening and corrected QP energies; (iii) solving the $e$-$h$ interaction in the BSE and calculating the optical absorption spectra. A good review can be found in [27]. This scheme has been applied to the calculations of absorption spectra of a wide range of materials and the simulation results agree with the experimental data quite well [30,44–47].

2.2.1 GW Approximation

The application of Green’s function techniques to electron-correlation problems is the basis of MBPT. The details of Green’s function approaches are discussed in the original paper by Hedin [31] and the review of Hedin and Lundqvist [32]. Here we just summarize the results directly relevant to the approximation of the self-energy operator of QPs.

The self-energy operator $\Sigma$ should in principle be calculated together with the Green’s function $G$ in a self-consistent procedure. They are related to each other...
as well as the screened Coulomb interaction $W$ and a vertex function $\Gamma$ by Hedin’s equations:

\[ \Sigma(1, 2) = i \int W(4, 1^+) G(1, 3) \Gamma(3, 2; 4)d(3)d(4); \]  
(2.37)

\[ W(1, 2) = \nu(1, 2) + \int W(4, 2) P(3, 4) \nu(1, 3)d(3)d(4); \]  
(2.38)

\[ P(1, 2) = -i \int G(1, 3) G(4, 1^+) \Gamma(3, 4; 2)d(3)d(4); \]  
(2.39)

\[ \Gamma(1, 2; 3) = \delta(1, 2) \delta(1, 3) + \int \frac{\delta \Sigma(1, 2)}{\delta G(4, 5)} G(4, 6) G(7, 5) \Gamma(6, 7; 3)d(4)d(5)d(6)d(7). \]  
(2.40)

Here the number notation is defined as $1 \equiv (r_1, t_1, \sigma_1)$, etc., $\sigma$ is the spin coordinate. $1^+$ means that $t \to t + \delta$ where $\delta$ is a positive infinitesimal. $\nu(1, 2)$ stands for the bare Coulomb interaction and $P$ denotes the irreducible polarizability.

The single-particle Green’s function is related to the spectral weight function $A(r, r'; E)$ in the following form:

\[ G(r, r'; E) = \int_C dE' \frac{A(r, r'; E')}{E - E'}. \]  
(2.41)

The contour $C$ runs infinitesimally above the real axis for $E' < \mu$ and below the real axis for $E' > \mu$ where $\mu$ is the chemical potential. The above formalism does not include the spin dependence explicitly. The spectral function is proportional to the imaginary part of the Green’s function:

\[ A(r, r'; E) = \frac{1}{\pi} |ImG(r, r'; E)|. \]  
(2.42)
By using the Lehmann representation in terms of QP energies $\varepsilon_s$ and amplitudes $f_s$, the Green’s function is written as:

$$
G(r, r'; E) = \sum_s \frac{f_s(r)f_s^*(r')}{E - [\varepsilon_s + i\eta \text{sgn}(\mu - \varepsilon_s)]},
$$

(2.43)

where the small imaginary part $\eta$ is needed for the convergence of the Fourier transform to the frequency space.

Then the Green’s function is related to solving the Dyson’s equation including the self-energy operator:

$$
\int \{ \delta(r - r')[\varepsilon_s - H_H(r)] - \Sigma(r, r'; \omega_s) \} f_s(r')dr' = 0,
$$

(2.44)

where $H_H$ contains the kinetic-energy operator, the external and the Hartree potential, $\Sigma$ includes the exchange and correlation interactions.

The above correlated relations between the Green’s function $G$ and the self-energy operator $\Sigma$ make it difficult to obtain exact solutions of the self-energy of QPs. However, as Hedin shows [32], the self-energy can be expanded as a perturbation series in terms of the dynamically screened Coulomb interaction (W) and the dressed Green’s function (G).

Hedin starts with the simplest approximation of the vertex function:

$$
\Gamma(1, 2; 3) = \delta(1, 2)\delta(1, 3).
$$

(2.45)
This is called the random-phase approximation (RPA) for the dielectric matrix. With RPA, the polarizability is given by

\[ P(1, 2) = -iG(1, 2^+)G(2, 1). \] (2.46)

Then the self-energy operator is simplified to the first order in \( W \):

\[ \Sigma(1, 2) = iG(1, 2)W(1^+, 2). \] (2.47)

The self-energy operator becomes the product of one-particle Green’s function and the screened Coulomb interaction term. This is called the GW approximation.

As we showed in the pseudopotential methods part, it is convenient to divide electrons of the solids into two parts: core electrons and valence electrons. Accordingly, the Green’s function and polarizability can be separated to \( G = G_c + G_v \) and \( P = P_c + P_v \), respectively. Then the electron self-energy operator can be written as

\[ \Sigma = i(G_c W + G_v W_v P_c W_v + G_v W_v). \] (2.48)

In realistic calculations, the first two terms are generally small and not sensitive to environment changes and excitations. So we only focus on the final term which is the self-energy of the valence electrons. In the following part, we will discuss the first-principles calculations of the electron self-energy and QP energies based on the framework in Ref. [29].
After the Fourier transform to energy, the GW approximation for $\Sigma$ becomes

$$\Sigma(r, r'; E) = i \int \frac{dE'}{2\pi} e^{-iE'} G(r, r'; E - E') W(r, r'; E'),$$

(2.49)

where $\delta = 0^+$.

Once the self-energy operator $\Sigma$ is known, the QP energies and wavefunctions can be obtained by solving the equation:

$$(K + V_{\text{ext}} + V_H) \psi_{nk}(r) + \int d\mathbf{r}' \Sigma(r, \mathbf{r}' ; E_{nk}) \psi_{nk}(\mathbf{r}') = E_{nk} \psi_{nk}(r),$$

(2.50)

where $K$ is the kinetic energy operator, $V_{\text{ext}}$ is the external potential due to the ions (nuclei + core electrons), $V_H$ is the average Coulomb (Hartree) potential due to valence electrons. Since $\Sigma$ is non-Hermitian, Eq. (2.50) generally gives complex eigenvalues $E_{nk}$; the real part of $E_{nk}$ is the QP energy and the imaginary part gives the lifetime of the QP.

The essential part of the QP problem is to obtain the self-energy operator $\Sigma$ by Eq. (2.49).

First, the screened interaction can be decomposed analogously to Eq. (2.41) as

$$W(r, r'; E) = v(r, r') + \int_0^{\infty} dE' \frac{2E'B(r, r'; E')}{E^2 - (E' - i\eta)^2},$$

(2.51)

where $\eta = 0^+$ and $B$ is the spectral function for the screened interaction, which is proportional to the imaginary part of $W$. 

46
As we can see in Eq. (2.43) and Eq. (2.51), the Green’s function $G$ and the screened Coulomb interaction $W$ both have poles in the energy domain. Therefore, with the Coulomb hole + Screened exchange (COHSEX) approximation, the real part of $\Sigma$ is divided into two parts:

$$Re \Sigma = \Sigma_{SEX} + \Sigma_{COH},$$

(2.52)

where $\Sigma_{SEX}$ is the screened exchange (SEX) term arising from the poles in the Green’s function, and $\Sigma_{COH}$ is the Coulomb-hole (COH) term arising from the poles in the screened interaction.

To further calculate the self-energy, approximations for $G$ and $W$ are necessary. The Green’s function can be approximately constructed with the resulting spectrum $\varepsilon_{nk}$ and wavefunctions $\phi_{nk}$ from DFT within LDA. The explicit expression of $G$ is

$$G(r, r'; E) = \sum_{n,k} \frac{\phi_{nk}(r) \phi^*_{nk}(r')}{E - \varepsilon_{nk} - i\delta_{nk}},$$

(2.53)

where $\delta_{nk} = 0^+$ for $\varepsilon_{nk} < \mu$, and $\delta_{nk} = 0^-$ for $\varepsilon_{nk} > \mu$, $\mu$ is the chemical potential.

Then the real part of the self-energy can be written as

$$\Sigma_{SEX}(r, r'; E) = -\sum_{n,k,occ} \phi_{nk}(r) \phi^*_{nk}(r') W(r, r'; E - \varepsilon_{nk});$$

(2.54)

$$\Sigma_{COH}(r, r'; E) = \sum_{n,k} \phi_{nk}(r) \phi^*_{nk}(r') P \int_{0}^{\infty} dE' \frac{B(r, r'; E')}{E - \varepsilon_{nk} - E'}. $$

(2.55)
From the above expressions, it’s clear that $\Sigma_{SEX}$ represents the dynamically screened-exchange interaction, and the $\Sigma_{COH}$ can be interpreted as the interaction of the QP with the induced potential due to the rearrangement of the electrons around the QP.

After the approximation of the Green’s function, we now have to deal with the screened Coulomb interaction $W$. The screened interaction can be described by the fully inverse dielectric matrix $\epsilon^{-1}(r, r'; \omega)$, where the relation is:

$$W(1, 2) = \int d(3) \epsilon^{-1}(1, 3) \upsilon(3, 2).$$

(2.56)

By the Fourier transform Eq. (2.56) to the energy and wave-vector space, the screened interaction is expressed as:

$$W_{GG'}(q, \omega) = \epsilon^{-1}_{GG'}(q, \omega) \upsilon(q + G'),$$

(2.57)

where $G$ is a reciprocal lattice vector and $q$ is a wave vector in the first Brillouin zone. The Fourier transform of the bare Coulomb interaction is $\upsilon(q) = 4\pi e^2/\Omega q^2$, in which $\Omega$ is the crystal volume.

While RPA is applied for the irreducible polarizability $P$ in Eq. (2.46), the dielectric matrix in the k-space is related to $P$ by

$$\epsilon_{GG'}(q, \omega) = \delta_{GG'} - \upsilon(q + G) P_{GG'}(q, \omega).$$

(2.58)
Then the calculation of the dynamic screening transfers to the calculation of polarizability. A convenient strategy is to first study the static screening and then extend it to the dynamical case.

According to Adler and Wiser’s work [65, 66], the static polarizability can be calculated from the spectrum $\varepsilon_{n,k}$ and wavefunctions $|n,k\rangle$ of DFT solutions within LDA. The relation is given by

$$P_{GG'}^0(q,\omega=0) = \sum_{n'k} \langle n,k|e^{-i(q+G)\cdot r}|n',k+q\rangle \langle n',k+q|e^{-i(q+G')\cdot r'}|n,k\rangle$$

$$\times \frac{f(\varepsilon_{n',k+q}) - f(\varepsilon_{n,k})}{\varepsilon_{n',k+q} - \varepsilon_{n,k}},$$

(2.59)

where $f(\varepsilon)$ is the fermion occupation factor.

Then the static dielectric matrix is obtained by Eq. (2.58), and the result can be extended to the frequency-dependent case with the general plasmon pole model (GPP) [72]. The GPP model is based on the observation that $\varepsilon^{-1}_{GG'}(q,\omega)$ is generally a peaked function in frequency $\omega$ and thus can be approximated by a single-pole function in $\omega$ [29].

Under GPP, the imaginary part of the dynamical inverse dielectric matrix is

$$Im\varepsilon^{-1}_{GG'}(q,\omega) = A_{GG'}(q)\{\delta(\omega - \tilde{\omega}_{GG'}(q)) - \delta(\omega + \tilde{\omega}_{GG'}(q))\},$$

(2.60)

and the real part is

$$Re\varepsilon^{-1}_{GG'}(q,\omega) = 1 + \frac{\Omega^2_{GG'}(q)}{\omega^2 - \tilde{\omega}^2_{GG'}(q)}.$$

(2.61)
where $\tilde{\omega}_{GG'}(q)$ is the mode frequency, $A_{GG'}(q)$ is the amplitude, and $\Omega_{GG'}(q)$ is the effective bare plasmas frequency.

By using the Kramer-Kronig relation between the real part and the imaginary part, and applying Johnso sum rules [67] to $Im\epsilon^{-1}_{GG'}(q, \omega)$, the three parameters can be obtained as:

$$
\Omega^2_{GG'}(q) = \omega^2_P \frac{(q + G) \cdot (q + G') \rho(G - G')}{|q + G|^2 \rho(0)}. \quad (2.62)
$$

$$
\tilde{\omega}^2_{GG'}(q) = \frac{\Omega^2_{GG'}(q)}{\delta_{GG'} - \epsilon^{-1}_{GG'}(q, \omega = 0)}. \quad (2.63)
$$

$$
A_{GG'}(q) = -\frac{\pi \Omega_{GG'}(q)}{2 \tilde{\omega}_{GG'}(q)}. \quad (2.64)
$$

Here $\rho$ is the charge density from the LDA result, and $\omega_P = \sqrt{4\pi n_e e^2/\mu}$ is the electron plasmon frequency.

The GPP model does not use any adjustable parameters and it generally provides a good description of the frequency dependence of $\epsilon^{-1}_{GG'}(q, \omega)$ in the important frequency range. Since the self-energy calculation involves an integration over the frequency domain, the GPP model usually works reasonably well for semiconductors and insulators [29].

In the first-principles calculations of Eq. (2.50), the QP wavefunction $\psi_{nk}(r)$ can be written as an expansion in a basis set consisting of the self-consistent wavefunctions $\phi_{nk}(r)$ from DFT within LDA:

$$
\psi_{nk}(r) = \sum_{n'} \alpha_{n,n'}(k) \phi_{n'k}(r). \quad (2.65)
$$
Accordingly, the self-energy can also be expanded in the $\phi_{nk}(\mathbf{r})$ basis using the dielectric function given by the GPP model. The formulation of the matrix elements of the self-energy is given by

$$
\langle n\mathbf{k}| \Sigma_{SEX}(\mathbf{r}, \mathbf{r}'; E)|n'\mathbf{k}'\rangle = -\sum_{n_1, q, G, G'}^{occ} \langle n\mathbf{k}| e^{i(\mathbf{q}+\mathbf{G})\cdot \mathbf{r}}|n_1, \mathbf{k} - \mathbf{q}\rangle \langle n_1, \mathbf{k} - \mathbf{q}| e^{-i(\mathbf{q}+\mathbf{G})\cdot \mathbf{r}'}|n', \mathbf{k}'\rangle \times [1 + \frac{\Omega_{GG'}^2(\mathbf{q})}{(E - \varepsilon_{n_1, k-q})^2 - \tilde{\omega}_{GG'}^2(\mathbf{q})}]v(\mathbf{q} + \mathbf{G}').
$$

(2.66)

$$
\langle n\mathbf{k}| \Sigma_{COH}(\mathbf{r}, \mathbf{r}'; E)|n'\mathbf{k}'\rangle = \sum_{n_1, q, G, G'} \langle n\mathbf{k}| e^{i(\mathbf{q}+\mathbf{G})\cdot \mathbf{r}}|n_1, \mathbf{k} - \mathbf{q}\rangle \langle n_1, \mathbf{k} - \mathbf{q}| e^{-i(\mathbf{q}+\mathbf{G})\cdot \mathbf{r}'}|n', \mathbf{k}'\rangle \times \frac{\Omega_{GG'}^2(\mathbf{q})}{2 \tilde{\omega}_{GG'}(\mathbf{q})[E - \varepsilon_{n_1, k-q} - \tilde{\omega}_{GG'}(\mathbf{q})]}v(\mathbf{q} + \mathbf{G}').
$$

(2.67)

Here, the state $|n\mathbf{k}\rangle = \phi_{nk}(\mathbf{r})$.

So far, the QP problem can be solved by combining the Eq. (2.50) together with Eqs. (2.65), (2.66) and (2.66).

In practical calculations of semiconductors and insulators, it is found that the non-diagonal terms of the self-energy only give small corrections to final solutions [29]. So only the calculation of the diagonal terms is necessary. Furthermore, the resulting QP wavefunctions are found to have better than 99.9% overlap with the corresponding LDA wavefunctions [29], indicating that $\phi_{nk}(\mathbf{r})$ is a very good approximation of $\psi_{nk}(\mathbf{r})$. 

51
Therefore, the QP energy can be approximated as

\[ E_{\text{qp}}^{nk} = \varepsilon_{nk}^{LDA} - \langle nk | V_{XC}^{LDA} | nk \rangle + \langle nk | \Sigma(E_{nk}) | nk \rangle. \]  

(2.68)

In Eq. (2.68) the self-energy operator must still be evaluated at the QP energy. Thus the QP energy should be solved iteratively. A practical way is to expand the self-energy matrix element to the first order in the energy around \( \varepsilon_{nk} \), and the QP energy is explicitly expressed as

\[ E_{\text{qp}}^{nk} = E_{nk}^0 + \frac{\Delta \Sigma_{nk}(\varepsilon_{nk})/\Delta E}{1 - \Delta \Sigma_{nk}(\varepsilon_{nk})/\Delta E}(E_{nk}^0 - \varepsilon_{nk}), \]  

(2.69)

where

\[ E_{nk}^0 = \varepsilon_{nk}^{LDA} - \langle nk | V_{XC}^{LDA} | nk \rangle + \langle nk | \Sigma(E_{nk}^{LDA}) | nk \rangle. \]  

(2.70)

The Eq. (2.69) involves a finite difference with \( \Delta E \) as the energy derivative evaluation. The typical value of \( \Delta E \) is chosen as 1 eV, and the results are quite insensitive to a reasonable choice of \( \Delta E \) as shown by Hybertsen and Louie [29]. However, for some other systems in which the QP energy is not far from the plasmon frequency, the self-consistent scheme may be needed.

The above framework of first-principles QP calculations with the GW approximation of self-energy provides a powerful tool to calculate trustworthy electronic band structure of solids. The GW calculations generally yield an excellent agreement between the experiment and the calculated band structure [29,35–38].
2.2.2 Bethe-Salpeter Equation (BSE) and Optical Absorption

As we have mentioned in the introduction part, the investigation of optical \(e-h\) excitations requires an effective two-body approach going beyond the single-particle picture. The two-body Green’s function have to be used and the \(e-h\) interaction has to be considered. This leads to the BSE. In the following part, we will provide the major steps about the BSE according to the Refs. [30,69,73].

First, the one-particle Green’s function is defined by the ground-state configuration \(|N, 0\rangle\) as:

\[
G_1(1, 2) = -i\langle N, 0|T(\psi(1)\psi^\dagger(2))|N, 0\rangle, \tag{2.71}
\]

where space, spin, and time coordinates are indicated and abbreviated by numbers, i.e., \((1) = (r_1, \sigma_1, t_1) = (x_1, t_1)\).

The two-particle Green’s function is defined as

\[
G_2(1, 2; 1', 2') = -\langle N, 0|T(\psi(2)\psi(1')\psi^\dagger(2')\psi^\dagger(1') )|N, 0\rangle. \tag{2.72}
\]

Then the \(e-h\) correlation function is defined as

\[
L(1, 2; 1', 2') = [G_2(1, 2; 1', 2') - G_1(1, 1')G_1(2, 2')]|_{\omega=0}. \tag{2.73}
\]

Here \(\omega\) is the external time-dependent perturbation.
Following Strinati [69], the evolution of the e-h correlation function satisfies the BSE in the form of

\[ L(1, 2; 1', 2') = L_0(1, 2; 1', 2') + \int L_0(1, 4; 1', 3)K(3, 5; 4, 6)L(6, 2; 5, 2')d(3)d(4)d(5)d(6), \]  
\[ \text{(2.74)} \]

where

\[ L_0(1, 2; 1', 2') = G_1(1, 2')G_1(2, 1'), \]  
\[ \text{(2.75)} \]

and \( K(3, 5; 4, 6) \) is the e-h interaction kernel. The details of \( K(3, 5; 4, 6) \) will be discussed later. Eq. (2.74) clearly shows that \( L_0(1, 2; 1', 2') \) corresponds to free e-h pairs with the interaction \( K \) being switched off.

In the context of optical excitations, only the difference between two sets of time variables is finally involved in Eq. (2.74) because of the time homogeneity in the absence of external fields. Therefore, we can Fourier transform the correlation function from the time spaces to one-dimensional energy (frequency) space.

Since the one-particle Green’s function can be expressed with the QP solutions (see Eq. (2.43)) from the GW calculations, \( L_0 \) can be written as

\[ L_0(1, 2; 1', 2'; \omega) = i \sum_{v,c} \left[ \frac{\psi_c(x_1)\psi_v^*(x_1')\psi_v(x_2)\psi_c^*(x_2')}{\omega - (E_c - E_v)} - \frac{\psi_v(x_1)\psi_c^*(x_1')\psi_c(x_2)\psi_v^*(x_2')}{\omega + (E_c - E_v)} \right], \]  
\[ \text{(2.76)} \]

where the indices, \( v \) and \( c \), run over the occupied hole states and empty electron states, respectively. For convenience and clarity, the imaginary infinitesimals in the denominators are not shown here.
Similarly, the correlation function \( L \) of Eq. (2.74) can be written with \( e-h \) excitation states amplitudes \( \chi_S(x, x') \) and corresponding excitation energies \( \Omega_S \) as:

\[
L(1, 2; 1', 2'; \omega) = i \sum_S \left[ \frac{\chi_S(x_1, x'_1) \chi_S^*(x'_2, x_2)}{\omega - \Omega_S} - \frac{\chi_S(x_2, x'_2) \chi_S^*(x'_1, x_1)}{\omega + \Omega_S} \right],
\]  

(2.77)

where

\[
\chi_S(x, x') = -\langle N, 0 \mid \psi_\downarrow(x') \psi(x) \mid N, S \rangle.
\]  

(2.78)

The \( e-h \) amplitudes can be further expressed in the basis given by the single-particle wavefunctions of the electron and hole states:

\[
\chi_S(x, x') = \sum_v \sum_c \left[ A_{vc}^S \psi_c(x) \psi_\downarrow^*(x') + B_{vc}^S \psi_\downarrow(x) \psi_c^*(x') \right].
\]  

(2.79)

With the help of Eqs. (2.76), (2.77), and (2.79), the BSE (Eq. (2.74)) becomes a generalized eigenvalue problem:

\[
(E_c - E_v) A_{vc}^S + \sum_{v'c'} K_{vc,v'c'}^{AA}(\Omega_S) A_{v'c'}^S + \sum_{v'c'} K_{vc,v'c'}^{AB}(\Omega_S) B_{v'c'}^S = \Omega_S A_{vc}^S; \tag{2.80a}
\]

\[
(E_c - E_v) B_{vc}^S + \sum_{v'c'} K_{vc,v'c'}^{BB}(\Omega_S) B_{v'c'}^S + \sum_{v'c'} K_{vc,v'c'}^{BA}(\Omega_S) A_{v'c'}^S = -\Omega_S B_{vc}^S, \tag{2.80b}
\]

where the matrix elements of the interaction kernel terms \( K \) are

\[
K_{vc,v'c'}^{AA}(\Omega_S) = i \int d(3)d(4)d(5)d(6) \psi_v(x_4) \psi_\uparrow^*(x_3) K(35; 46; \Omega_S) \psi_\downarrow^*(x_5) \psi_c(x_6); \tag{2.81}
\]
\[ K^{AB}_{vc,v'}(\Omega_S) = i \int d(3)d(4)d(5)d(6)\psi_v(x_4)\psi^*_v(x_3)K(35;46;\Omega_S)\psi^*_{v'}(x_6)\psi_{v'}(x_5). \] (2.82)

and the expressions are similar for \( K^{BA} \) and \( K^{BB} \).

In realistic calculations, it is found that the off-diagonal blocks of \( K^{AB} \) and \( K^{BA} \) are small and have nearly no effect on the excitation energies [30, 46]. Therefore, \( A^{S}_{vc} \) and \( B^{S}_{vc} \) are decoupled when we set \( K^{AB} = K^{BA} = 0 \), and the two equations about \( A^{S}_{vc} \) and \( B^{S}_{vc} \) yield exactly the same excitations besides a sign difference for the excitation energies. Therefore, we only need to look at the below eigenvalue problem:

\[ (E_c - E_v)A^{S}_{vc} + \sum_{v'c'} K^{AA}_{vc,v'c'}(\Omega_S)A^{S}_{v'c'} = \Omega_S A^{S}_{vc}. \] (2.83)

This is equivalent to the Tamm-Dancoff approximation [47] that expands the excited states in e-h pair configuration as

\[ |N,S\rangle = \sum_{v} \sum_{c} A^{S}_{vc} \hat{a}^\dagger_v \hat{b}^\dagger_c |N,0\rangle =: \sum_{v} \sum_{c} A^{S}_{vc} |vc\rangle, \] (2.84)

where \( \hat{a}^\dagger_v \) and \( \hat{b}^\dagger_c \) create a hole or an electron, respectively, to the many-body ground state \( |N,0\rangle \), and \( |vc\rangle \) are the uncorrelated QP wavefunctions of electron states and hole states.
To solve Eq. (2.83), the e-h interaction kernel \( K \) has to been calculated. Here we have to specify this interaction kernel. From the two-body Green’s function theory, \( K \) is given by the functional derivative

\[
K(35; 46) = \frac{\delta[V_{\text{coul}}(3)\delta(3, 4) + \Sigma(3, 4)]}{\delta G_1(6, 5)}.
\]

By employing the GW approximation for the self-energy \( \Sigma \) and assuming that the derivative of the screened interaction \( W \) with respect to \( G_1 \) can be neglected, the interaction kernel breaks into two parts:

\[
K(35; 46) = -i\delta(3, 4)\delta(5^-, 6)v(3, 6) + i\delta(3, 6)\delta(4, 5)W(3^+, 4)
\]

\[
=: K^x(35; 46) + K^d(35; 46).
\]

The exchange term \( K^x \) results from the bare Coulomb potential while the direct interaction term \( K^d \) is from the screened-exchange self-energy in Eq. (2.85). The direct interaction term \( K^d(35; 46) \) includes the attractive nature of the e-h interaction and is responsible for the formation of bound e-h states \( \text{i.e.}, \) excitons. The exchange interaction term \( K^x(35; 46) \) is generally positive and contributes the splitting between spin-singlet and spin-triplet excitations.
The matrix elements of the interaction kernel in Eq. (2.83) are expressed as

$$\langle vc | K^{AA,d} (Ω_S) | v' c' \rangle = \int dx dx' \psi^*_c (x) \psi_d (x) \psi_d (x') \psi^*_c (x') \frac{i}{2\pi} \int d\omega e^{-i\omega \Omega_S} W(r, r'; \omega) \times \left[ \frac{1}{Ω_S - \omega - (E_{QP}^d - E_{QP}^v) + i0^+} + \frac{1}{Ω_S + \omega - (E_{QP}^d - E_{QP}^v) + i0^+} \right].$$

(2.87)

and

$$\langle vc | K^{AA,x} (Ω_S) | v' c' \rangle = \int dx dx' \psi^*_c (x) \psi_d (x) \psi_d (x') \psi^*_c (x') \times \omega^2 \left[ \frac{1}{Ω_S - \omega - (E_{QP}^d - E_{QP}^v) + i0^+} + \frac{1}{Ω_S + \omega - (E_{QP}^d - E_{QP}^v) + i0^+} \right].$$

(2.88)

By using the same plasmon-pole model employed in the previous GW method, the screened interaction can be written

$$W(x, x', \omega) = \sum_l W_l(x, x') \frac{\omega_l}{\omega - \omega_l + i0^+} - \frac{1}{\omega + \omega_l - i0^+},$$

(2.89)

where the index $l$ stands for the plasmon mode, $\omega_l$ is the plasmon frequency of the plasmon model $l$, and $W_l$ is the corresponding spatial behavior.

With the help of Eq. (2.89), the matrix element of the direct term $K^d$ becomes

$$\langle vc | K^{AA,d} (Ω_S) | v' c' \rangle = - \sum_l \int dx dx' \psi^*_c (x) \psi_d (x) \psi_d (x') \psi^*_c (x') W_l(r, r') \times \frac{\omega_l}{\omega_l - (Ω_S - (E_{QP}^d - E_{QP}^v))} \left[ \frac{1}{Ω_S - (E_{QP}^d - E_{QP}^v)} + \frac{1}{Ω_S + (E_{QP}^d - E_{QP}^v)} \right].$$

(2.90)
The interaction kernel in Eq. (2.90) depends on the e-h excitation energy $\Omega_S$. In many cases (e.g., in most semiconductor crystals) when $\omega_l \gg \Omega_S - (E_{QP}^c - E_{QP}^v)$, the exciton binding energies are far away from the plasmon frequency that controls the dynamics of the screening. Eq. (2.90) can be simplified with the static approximation to the form of

$$\langle vc | K^{AA,d} | v'c' \rangle = - \int dxdx' \psi_c^*(x) \psi_c(x) \psi_v(x') \psi_{v'}^*(x') W(r, r', \omega = 0). \quad (2.91)$$

However, in other cases when the exciton binding energies are of the same order of magnitude as the characteristic plasmon frequency, the static approximation is no longer valid and it requires an iterative procedure to calculate the interaction kernel and solve the BSE.

By now, we have specified all terms in BSE calculations. When considering the optical absorption, we have to include the spin in the Hamiltonian. We only consider the situation that the spin-orbit interaction is negligible (i.e., as compared to the e-h interaction), then the single-particle states ($v$ and $c$) can be classified as spin-up states ($v \uparrow$ and $c \uparrow$) and spin-down states ($v \downarrow$ and $c \downarrow$). Note that for occupied states $v$, this spin corresponds to the electron that originally occupies the state; the spin of the corresponding hole would be the negative. The Hilbert space of the e-h pairs ($vc$) now consists of four subspaces: $v \uparrow c \uparrow$, $v \uparrow c \downarrow$, $v \downarrow c \uparrow$ and $v \downarrow c \downarrow$. Between these subspaces, most matrix elements of Eqs. (2.87) and (2.88) are zero.
Let $D \equiv (E_c - E_v)$, then in the notation of these subspaces (in an order of $v \uparrow c \uparrow$, $v \uparrow c \downarrow$, $v \downarrow c \uparrow$ and $v \downarrow c \downarrow$), the BSE Hamiltonian obtains the following form:

$$H^{ch} = \begin{pmatrix}
D + K^d + K^x & 0 & 0 & K^x \\
0 & D + K^d & 0 & 0 \\
0 & 0 & D + K^d & 0 \\
K^x & 0 & 0 & D + K^d + K^x
\end{pmatrix} \quad (2.92)$$

This Hamiltonian decouples into two classes: a spin-triplet class with the Hamiltonian eigenvalue $D + K^d$ and the solutions consisting of the subspaces $v \uparrow c \downarrow$, $v \downarrow c \uparrow$, and $\frac{1}{\sqrt{2}}(v \uparrow c \uparrow + v \downarrow c \downarrow)$; a spin-singlet class with Hamiltonian eigenvalue $D + K^d + K^x$ and the solutions consisting of the subspace $\frac{1}{\sqrt{2}}(v \uparrow c \uparrow - v \downarrow c \downarrow)$. Therefore, the BSE can be solved for singlet and triplet cases separately. The splitting between spin-singlet and spin-triplet excitations is controlled by $K^x$, the exchange term of e-h interaction kernel.

Because of the selection rule, only the singlet state contributes to the optical absorption. Thus the spin degrees of freedom have been completely eliminated from the remaining problem by focusing only on the singlet state. The interaction of an external light field with excitations in the system is described by the imaginary part of macroscopic transverse dielectric function of the system,

$$\epsilon_2(\omega) = \frac{16\pi e^2}{\omega^2} \sum_s |\vec{\lambda} \cdot \langle 0 | \vec{v} | S \rangle|^2 \delta(\omega - \Omega_s). \quad (2.93)$$
Here $\vec{\lambda} = \vec{A}/|\vec{A}|$ is the polarization vector of the light and $\vec{v} = \frac{i}{\hbar}[H, \vec{r}]$ is the single-particle velocity operator.

Eq. (2.93) requires the $e$-$h$ excitation energies and wavefunctions as input, the optical transition matrix elements $\langle 0|\vec{v}|S \rangle$ are crucial to the calculation of optical absorption. Without the $e$-$h$ interaction, the excitations are given by those vertical transitions between independent hole and electron states, and Eq. (2.93) reduces to the form

$$
e_2^0(\omega) = \frac{16\pi\epsilon^2}{\omega^2} \sum_{v,c} |\vec{A} \cdot \langle v|\vec{v}|c \rangle|^2 \delta(\omega - (E_c - E_v)), \quad (2.94)$$

where $v$ denotes valence states and $c$ denotes conduction states.

By using Eq. (2.84), the optical transition matrix elements can be calculated through a coherent sum of the contributing vertical transition matrix elements including the coupling coefficients $A_{vc}^S$:

$$\langle 0|\vec{v}|S \rangle = \sum_v^{\text{hole}} \sum_c^{\text{elec}} A_{vc}^S \langle v|\vec{v}|c \rangle. \quad (2.95)$$

In summary, we draw a flow chart of the first-principles MBPT calculation, as shown in Figure 2.3. The calculation starts from DFT, solving the Kohn-Sham equations which includes the electron exchange and correlations by using approximated functionals such as LDA. The solutions of Kohn-Sham equations are then used to construct the one-particle Green’s function and static dielectric matrix at the zeroth order. Using the GPP model, the frequency-dependent dielectric matrix is constructed and the dynamic screening is obtained. After the RPA, the self-energy is calculated.
Figure 2.3. The flow chart of MBPT for calculating the QP band structure and optical absorption spectra.
by the GW approximation as the product of the one-particle Green’s function and the screened Coulomb interaction. Then the QP equations are solved with the self-energy replacing the LDA approximated functionals. Since the self-energy depends on the QP energies, it can first use Kohn-Sham energies as the zeroth-order input, and then solve the QP energies by the self-consistent scheme. For the optical excitations of $e$-$h$ pairs, the QP energies and wavefunctions are used to calculate the interaction kernel with the dynamic screening. In realistic calculations, QP wavefunctions are approximated by the Kohn-Sham wavefunctions since they are very close to each other for most materials. With the interaction kernel and the Tamm-Dancoff approximation, the BSE is solved and the excitation energies and the $e$-$h$ amplitudes are obtained. Finally, based on the BSE solutions, the optical absorption spectrum can be evaluated through a coherent summation of the contributing vertical transition matrix elements including the coupling coefficients of $e$-$h$ amplitudes.

2.3 First-Principles Phonon Calculations

A phonon is a quantum-mechanical description of an elementary vibrational motion, in which atoms or molecules of lattices uniformly oscillates at a single frequency. It is a QP concept of an excited state of the collective atomic vibrations. Phonons play a major role in many physical properties of solids, such as thermal conductivity and electrical conductivity. The study of phonons is an essential part of condensed matter physics.
In this section, we start with a simple derivation of lattice vibrations in a harmonic crystal. Then we discuss the linear response theory for the first-principles phonon calculations.

2.3.1 Lattice Vibrational Modes in Harmonic Crystals

In a crystal, the equilibrium position of an ion is $\mathbf{R}$. Assuming applying a small perturbation to the ion induces a displacement $\mathbf{u}(\mathbf{R})$, the position of the ion becomes $\mathbf{r}(\mathbf{R}) = \mathbf{R} + \mathbf{u}(\mathbf{R})$. For simplicity, we consider only one atom per unit cell in the following derivations. For the case that there are more than one atom per unit cell, it is easy to extend the derivations by including more ions with positions defined in the same way as $\mathbf{r}_j(\mathbf{R}) = \mathbf{R} + \mathbf{d}_j + \mathbf{u}_j(\mathbf{R})$, where the index $j$ runs over all the atoms in the unit cell.

The total static energy depends on the positions of ions as:

$$E = E(\mathbf{R} + \mathbf{u}(\mathbf{R})). \quad (2.96)$$

Using the Taylor expansion of total static energy of the crystal with respect to the small perturbation of ion positions from their equilibrium and only keeping the terms up to the second order due to the harmonic approximation, we obtain the form

$$E \approx E|_{\mathbf{u}=0} + \sum_{\mathbf{R}} \mathbf{u}(\mathbf{R}) \cdot \left[ \nabla_u E \right]|_{\mathbf{u}=0} + \frac{1}{2} \sum_{\mathbf{R},\mathbf{R}'} \mathbf{u}(\mathbf{R}) \cdot D(\mathbf{R},\mathbf{R}') \cdot \mathbf{u}(\mathbf{R}'), \quad (2.97)$$
where

\begin{equation}
D_{\mu\nu}(R, R') = \left. \frac{\partial^2 E}{\partial u_\mu(R) \partial u_\nu(R')} \right|_{u=0}.
\end{equation}

(2.98)

Here \(\mu\) and \(\nu\) denote the displacement directions.

Because the crystal is perturbed from the equilibrium, the first term in Eq. (2.97) is just a constant and the second term turns out to be zero, the only interesting term is the third term which is also called the harmonic energy

\begin{equation}
\Delta E \approx E^{\text{harmonic}} = \frac{1}{2} \sum_{R, R'} \sum_{\mu, \nu} u_\mu(R) D_{\mu\nu}(R, R') u_\nu(R').
\end{equation}

(2.99)

In the harmonic approximation, the tensor \(D_{\mu\nu}(R, R')\) describes the force acting on the ion at \(R\) along the \(\mu\) direction induced by the unit displacement of the ion at \(R'\) along the \(\nu\) direction, thus \(D_{\mu\nu}(R, R')\) is named the force matrix.

Because of the translational symmetry of the crystal, the force matrix \(D_{\mu\nu}(R, R')\) depends only on the difference between \(R\) and \(R'\), which is

\begin{equation}
D_{\mu\nu}(R, R') = \left. \frac{\partial^2 E}{\partial u_\mu(R) \partial u_\nu(R')} \right|_{u=0} = \left. \frac{\partial^2 E}{\partial u_\mu(R - R') \partial u_\nu(0)} \right|_{u=0} = D_{\mu\nu}(R - R').
\end{equation}

(2.100)

The force matrix obeys two symmetry rules. The first one is that the force matrix is invariant under the inversion operation,

\begin{equation}
D_{\mu\nu}(R - R') = D_{\mu\nu}(R' - R).
\end{equation}

(2.101)
This rule is based on the inversion symmetry of Bravais lattices.

The second one is that the force matrix obeys the sum rule:

\[ \sum_{\mathbf{R}} D_{\mu\nu}(\mathbf{R}) = 0. \]  \hspace{1cm} (2.102)

This sum rule implies that if all ions are shifted by the same displacement, meaning that the whole lattices shift, there is no force acting on the ions.

With the definition of the force matrix, the dynamics of ions with the mass \( M \) is described by the equation:

\[ M \ddot{u}_\mu(\mathbf{R}) = - \sum_{\mathbf{R}'} \sum_{\nu} D_{\mu\nu}(\mathbf{R} - \mathbf{R}') u_\nu(\mathbf{R}'). \]  \hspace{1cm} (2.103)

This is a harmonic oscillation and the solution of Eq. (2.103) has the form of

\[ u(\mathbf{R}, t) = c e^{i (\mathbf{k} \cdot \mathbf{R} - \omega t)}, \]  \hspace{1cm} (2.104)

where \( c \) is the polarization unit vector, and \( \mathbf{k} \) is a vector in the reciprocal space. In order to solve \( c \) and \( \omega \), we take the Fourier transform of the force matrix and obtain its reciprocal space correspondent, which is the dynamical matrix:

\[ \mathbf{D}(\mathbf{k}) = \sum_{\mathbf{R}} \mathbf{D}((\mathbf{R}) e^{i \mathbf{k} \cdot \mathbf{R}}). \]  \hspace{1cm} (2.105)
Combining Eqs. (2.103), (2.104), and (2.105), the dynamical equation in the reciprocal space is derived as

\[ M\omega^2 \mathbf{c} = \mathbf{D}(\mathbf{k}) \cdot \mathbf{c}. \]  

(2.106)

Clearly, Eq. (2.106) is an eigenvalue problem. By diagonalizing the dynamical matrix \( \mathbf{D}(\mathbf{k}) \), the eigenfrequencies \( \omega \) and the eigenmodes \( \mathbf{c} \) of the lattice vibrations can be obtained. It has to be pointed out that the vibration modes are associated with the wave-vector \( \mathbf{k} \) in the reciprocal space. For different choice of \( \mathbf{k} \), \( \mathbf{D}(\mathbf{k}) \) are different. Therefore, the eigenfrequencies \( \omega \) and the eigenmodes \( \mathbf{c} \) are \( \mathbf{k} \) dependent. The dispersion relation \( \omega(\mathbf{k}) \) gives the structure of normal modes and provides a full description of vibrational characteristics of the crystal.

In quantum mechanics, the collective excitation of crystal vibrations in a single frequency is called a phonon. Phonon is a “quasiparticle” (boson) picture of collective vibrations and its energy is quantized with \( \hbar \omega(\mathbf{k}) \). Within the phonon description, the Hamiltonian of the harmonic crystal can be written as

\[ H = \sum_{\mathbf{k}} \hbar \omega(\mathbf{k})(a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2}), \]  

(2.107)

where the creation operator \( a_{\mathbf{k}} \) and the annihilation operator \( a^\dagger_{\mathbf{k}} \) are defined as

\[ a_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{-i\mathbf{k} \cdot \mathbf{R}} c(\mathbf{k}) : \left[ \sqrt{\frac{M\omega(\mathbf{k})}{2\hbar}} \mathbf{u}(\mathbf{R}) + i \sqrt{\frac{1}{2\hbar M\omega(\mathbf{k})}} \mathbf{P}(\mathbf{R}) \right], \]  

(2.108)
and

$$a_k^\dagger = \frac{1}{\sqrt{N}} \sum_R e^{-i\mathbf{k} \cdot \mathbf{R}} c(\mathbf{k}) \cdot \left[ \sqrt{\frac{M \omega(\mathbf{k})}{2\hbar}} \mathbf{u}(\mathbf{R}) - i \sqrt{\frac{1}{2\hbar M \omega(\mathbf{k})}} \mathbf{P}(\mathbf{R}) \right].$$

(2.109)

Here $\mathbf{u}(\mathbf{R})$ is the real-space eigenmode operator and $\mathbf{P}(\mathbf{R})$ is the momentum operator of the eigenmode.

From the above derivations, the most essential part of studying lattice vibrations is to calculate the dynamical matrix or the force matrix. In first-principles phonon calculations, there are mainly two approaches to calculate them.

The first one is the small-displacement method which is also referred as the “direct” method [48, 74, 75]. The calculation is based on the fact that the force matrix expresses the proportionality between forces and displacements, and the relation is linear when the displacements are small enough. In this case, the force matrix $D_{\mu\nu}(\mathbf{R}, \mathbf{R}')$ describes the force acting on the ion at $\mathbf{R}$ along the $\mu$ direction induced by the unit displacement of the ion at $\mathbf{R}'$ along the $\nu$ direction. The forces acting on the ions can be calculated according to the Feynman-Hellman theorem [76]. Therefore, what we need to do is to displace a single atom in the crystal along a Cartesian direction, with all the other atoms being fixed at their equilibrium positions; the resulting forces on all the atoms give directly the elements of the force constant matrix. If this procedure is repeated for every single atom along every direction, all the elements of the force-constant matrix can be obtained, and then the phonon frequencies can be calculated from the above derivations.
In practice, it’s found that the force acting on ions decays rapidly with the increasing separation from the perturbed ion. Therefore, a cutoff distance can be specified. This small-displacement method is rather straightforward computationally, and it is easy to be implemented into the total-energy code. Additionally, this method includes both linear and nonlinear effects directly. However, the convergence of the force matrix can be slow in the periodic materials of semiconductors and insulators, in which the forces induced by perturbations are relative long-range and thus a large cutoff distance is required to ensure the convergence. The method is also computationally expensive for a system containing a large number of atoms per unit cell.

The second one is the linear response method [50–54, 77]. The main idea is from the Hellmann-Feynman theorem that a linear order variation of the electron density upon a perturbation applied to the crystal is responsible for a variation of the total energy up to the second order of the perturbation [54, 76]. If the system is perturbed in a phonon wave with a wave-vector \( k \), the corresponding change in electron density to the first order can be obtained by using only unperturbed wavefunctions, and it can further be used to determine the dynamical matrix (DM) at this wave vector \( k \) [54]. Therefore, only the ground-state calculations of crystal are required, and this procedure can be done for any arbitrary wave vector, without the need for the construction of a supercell.

As in its name, the linear response method only considers the linear effect of the perturbation. It needs extra steps to calculate higher-order terms for the nonlinear effect. The merits of this method reside particularly in the phonon calculations of
seminconductors and insulators. The linear-response method does not need to construct a supercell and it overcomes the slow convergence problem in the limit of zero wave-vector, which significantly increases the efficiency of the computation.

In our first-principles phonon calculations of reduced dimensional materials, we mainly use the linear response method. Therefore, I will introduce more details of the linear response theory in the next subsection.

2.3.2 Linear Response Theory

The linear response theory provides a general framework for analyzing the dynamical properties of a physical system close to an equilibrium or a steady state. The equilibrium system responds to a perturbation such as an external field or fluctuations of the control parameters, which may depend on time. When the perturbation is small, the induced change or the response of the system is assumed to be a linear function of the perturbation approximately. The linear response function provides valuable information of the system, such as specific heat, magnetic and electric susceptibility, etc. The linear response theory was first clearly formulated by Kubo in 1957 [78], and was later applied to many studies including phonon calculations.

The early work of the linear response method involves calculating the dielectric matrix and polarizability [51], which describes the linear response of the electronic charge density to the external perturbation. The dielectric matrix can be used for studying many interesting properties other than phonons, but it requires the inverse operation of a large-size matrix, making the computation quite heavy. Later, the
density-functional perturbation theory (DFPT) for phonon calculations [54,77] was
developed, and an efficient self-consistent approach to the linear response [52,53] was
proposed. Nowadays the first-principles linear response method for phonon calcula-
tions has been implemented into the package Quantum Espresso [79].

The key ideas of this method are summarized here. Let us assume that the external
potential depends on a perturbation parameter $\lambda$:

$$V_\lambda(r) \simeq V_0(r) + \frac{\partial V(r)}{\partial \lambda} \lambda + \frac{1}{2} \frac{\partial^2 V(r)}{\partial \lambda^2} \lambda^2 + \ldots. \quad (2.110)$$

Here all derivatives are calculated at $\lambda = 0$. For convenience, we do the same thing
for all the following derivatives unless we specify explicitly.

The corresponding electronic charge density with respect to the perturbation can
also be expanded as

$$n_\lambda(r) \simeq n_0(r) + \frac{\partial n(r)}{\partial \lambda} \lambda + \frac{1}{2} \frac{\partial^2 n(r)}{\partial \lambda^2} \lambda^2 + \ldots, \quad (2.111)$$

where $n_0$ is the ground-state charge density when $\lambda = 0$.

By the Hellmann-Feynman theorem, the first-order derivative $\frac{\partial E_\lambda}{\partial \lambda}$ does not depend
on any derivative of $n(r)$,

$$\frac{\partial E_\lambda}{\partial \lambda} = \int n_\lambda(r) \frac{\partial V_\lambda(r)}{\partial \lambda} \, dr. \quad (2.112)$$
Then, the total energy of the system under the perturbation can be written as

\[ E_\lambda \simeq E_0 + \lambda \frac{\partial E}{\partial \lambda} + \frac{1}{2} \lambda^2 \frac{\partial^2 E}{\partial \lambda^2} + \ldots. \]

\[ = E_0 + \lambda \int n_0(\mathbf{r}) \frac{\partial V_\lambda(\mathbf{r})}{\partial \lambda} \, d\mathbf{r} + \frac{1}{2} \lambda^2 \int \left[ \frac{\partial n_\lambda(\mathbf{r})}{\partial \lambda} \frac{\partial V_\lambda(\mathbf{r})}{\partial \lambda} + n_0(\mathbf{r}) \frac{\partial^2 V_\lambda(\mathbf{r})}{\partial \lambda^2} \right] d\mathbf{r} + \ldots. \]

(2.113)

For the case of lattice vibrations in a crystal, the parameter \( \lambda \) is regarded as the displacement of ions from their equilibrium positions. From the Hohenberg-Kohn theorem, we have \( \int n_0(\mathbf{r}) \frac{\partial V_\lambda(\mathbf{r})}{\partial \lambda} \, d\mathbf{r} = 0 \). Then we obtain a similar equation to Eq. (2.99):

\[ \Delta E(\lambda) \simeq \frac{1}{2} \lambda^2 \int \left[ \frac{\partial n_\lambda(\mathbf{r})}{\partial \lambda} \frac{\partial V_\lambda(\mathbf{r})}{\partial \lambda} + n_0(\mathbf{r}) \frac{\partial^2 V_\lambda(\mathbf{r})}{\partial \lambda^2} \right] d\mathbf{r}, \]

(2.114)

\( \Delta E(\lambda) \) is the harmonic energy under the perturbation \( \lambda \). It consists of two parts: one is from the electron-density response, the other is from the ionic energy change. Therefore, we define the force matrix as

\[ C^{\alpha\beta}_{\mu\nu}(\mathbf{R} - \mathbf{R'}) = C^{\alpha\beta,\text{elec}}_{\mu\nu}(\mathbf{R} - \mathbf{R'}) + C^{\alpha\beta,\text{ion}}_{\mu\nu}(\mathbf{R} - \mathbf{R'}). \]

(2.115)

Different from the definition of the force matrix in Eq. (2.98), here we include more than one atom per unit cell; \( \alpha \) and \( \beta \) are the indexes specifying the atoms inside a unit cell.

In Eq. (2.115), the ion part \( C^{\alpha\beta,\text{ion}}_{\mu\nu} \) is the second derivative of the Ewald energy [80], whose expression is in Eq. (2.36). The electronic part is the most difficult part.
to be evaluated since it is relative to the first-order derivative of the charge density. Using ion displacements as the perturbation, the electronic part of the force matrix has the form

$$C_{\mu \nu}^{\alpha \beta, \text{elec}}(\mathbf{R} - \mathbf{R}') = \int \left[ \frac{\partial n(\mathbf{r})}{\partial u_\mu^\alpha(\mathbf{R})} \frac{\partial V_{\text{ion}}(\mathbf{r})}{\partial u_\nu^\beta(\mathbf{R}')} + n_0(\mathbf{r}) \frac{\partial^2 V_{\text{ion}}(\mathbf{r})}{\partial u_\mu^\alpha(\mathbf{R})\partial u_\nu^\beta(\mathbf{R}')} \right] d\mathbf{r}, \tag{2.116}$$

where $V_{\text{ion}}$ is the ionic potential. In realistic calculations of a crystal, the ion potential is the sum of the pseudopotentials from all ions.

In the reciprocal space, the matrix above is represented as

$$C_{\mu \nu}^{\alpha \beta, \text{elec}}(\mathbf{q}) = \int \left[ \frac{\partial n(\mathbf{r})}{\partial u_\mu^\alpha(\mathbf{q})} \frac{\partial V_{\text{ion}}(\mathbf{r})}{\partial u_\nu^\beta(\mathbf{q})} + n_0(\mathbf{r}) \frac{\partial^2 V_{\text{ion}}(\mathbf{r})}{\partial u_\mu^\alpha(\mathbf{q})\partial u_\nu^\beta(\mathbf{q})} \right] d\mathbf{r}, \tag{2.117}$$

where $u_\mu^\alpha(\mathbf{q})$ is the Fourier transform of the displacement of ions at wave vector $\mathbf{q}$. In practical calculations, the displacement of ions is chosen according to the vibrational modes.

Finally, the dynamical matrix is defined as

$$D_{\mu \nu}^{\alpha \beta}(\mathbf{q}) = \frac{C_{\mu \nu}^{\alpha \beta}(\mathbf{q})}{\sqrt{M_\alpha M_\beta}}, \tag{2.118}$$

where $M_\alpha$ and $M_\beta$ are the mass of ion $\alpha$ and ion $\beta$, respectively.

Since the derivatives of $V_{\text{ion}}$ are easy to calculate, the key point for obtaining the dynamical matrix is evaluating $\frac{\partial n(\mathbf{r})}{\partial u_\mu^\alpha(\mathbf{q})}$, which is the linear response of the electron density according to the lattice perturbations. With the help of the DFT method,
\( \frac{\partial n(r)}{\partial \omega^q(q)} \) can be valued by an efficient self-consistent approach developed by Baroni, et al [52, 53, 77].

Within the single-particle picture for electrons, the effective potential of the particle is expressed as

\[
V_{SCF}(r) = V_{ion}(r) + \int \frac{n(r')}{|r - r'|} dr' + \mu_{xc}(n(r)). \tag{2.119}
\]

The perturbation of ions induces the ion potential change \( \Delta V_{ion} \), which results in a new single-particle effective potential \( V_{SCF} + \Delta V_{SCF} \). Under LDA, the Kohn-Sham equations have a form similar to the Schrödinger equation. Using the first-order perturbation theory, the linear term of the change of the density can be derived as

\[
\Delta n(q + G) = \frac{4}{N\Omega} \sum_k \sum_{c,v} \frac{\langle \psi_{v,k}|e^{-i(q+G)\cdot r}\psi_{c,k+q}\rangle \langle \psi_{c,k+q}|\Delta V_{SCF}|\psi_{v,k}\rangle}{\varepsilon_{v,k} - \varepsilon_{v,k+q}}, \tag{2.120}
\]

where \( \Delta n(q + G) \) is the Fourier transform of \( \Delta n(r) \). \(|\psi\rangle\) and \( \varepsilon \) are the wavefunctions and eigenvalues obtained from LDA calculations without the perturbation. The indices \( v \) and \( c \) denote the valence and conduction bands, respectively. The wave vector \( k \) is sampled from k-grids in the first Brillouin zone and \( G \) is the reciprocal lattice vector. \( \Omega \) is the volume of one unit cell.

By including the density change in Eq. (2.120) to Eq. (2.119), the change of the effective potential is given by

\[
\Delta V_{SCF} = \Delta V_{ion}(r) + \int \frac{\Delta n(r')}{|r - r'|} dr' + \Delta n(r) \frac{d\mu_{xc}}{dn}. \tag{2.121}
\]
Eqs. (2.120) and (2.121) consist of a self-consistent loop. The self-consistent procedure only requires the ground-state Kohn-Sham wavefunctions without perturbations, and the calculation effort for each wave vector $\mathbf{q}$ is equivalent. Therefore, this self-consistent approach is very efficient in simulations and the convergence is fast.

Once we obtain $\frac{\partial n(r)}{\partial \mathbf{q}}$, the dynamic matrix can be calculated by this linear response method. Finally the phonon frequencies and vibrational modes can be obtained by the matrix diagonalization.

It must be pointed out that Eq. (2.120) requires calculations of a large number of conduction bands, which hinder the computation efficiency. A new approach using the Green’s function method is developed to project the state $V_{SCF}|\psi_c\rangle$ into the valence band space, improving the efficiency of the simulations. The details of this Green’s function approach can be found in the references [52,53].
3. BAND OFFSETS ENGINEERING IN SI/GE CORE-SHELL NANOWIRES

3.1 Introduction of Heterojunctions and Core-Shell Nanowires

A heterojunction is an interface between two dissimilar crystalline semiconductors, which, more generally, could be any two solid-state materials. The most important feature of a heterojunction is that the electronic properties depend on spatial regions since the band gaps of the participating semiconductors are different. Thus there will be discontinuities in both the conduction and valence bands. These band discontinuities result in many useful electronic, transport, and optical properties of heterojunctions. In this sense, studying the band alignments and corresponding band offsets is of the central interest for heterojunctions and their applications.

Depending on how the band edges are aligned in the interface, heterojunctions are categorized into three types, which are: the straddled or type I lineup, the staggered or type II lineup, and the broken or type III lineup, as shown in Fig. 3.1. These band alignments provide a precious degree of freedom for tuning transport and electric properties.
Heterojunctions are of fundamental importance in condensed matter physics and are also elementary building blocks of microelectronics including semiconductor lasers, solid-state lighting, and transistors [81, 82]. As an example, Fig. 3.2 illustrates the physical mechanism of heterojunctions with a type II band alignment for photovoltaic (PV) solar cells application. When an electron is excited to the conduction band by the incident photon in the semiconductor B, it will leave a positively-charged hole in the valence band. These electrons and holes will attract each other to form the \( e-h \) pairs (excitons). When these excited excitons diffuse to the interface of the heterojunction, the electron will be drawn to the semiconductor A while the hole will stay in the semiconductor B because of the type II band alignment. As a result, the electron and hole pairs are separated, producing an output current. This mechanism shows that the efficiency of solar cells is decided by the band gaps of the involved semiconductors and, more importantly, the band offsets of the heterojunction.

Recently nanoscale radial heterojunctions in one-dimensional (1-D) core-shell nanowires (NWs) have been fabricated and have shown promising electrical and op-
Figure 3.2. Schematics of a heterojunction working for the photovoltaics solar cell. $\Delta E_c$ and $\Delta E_v$ denote the conduction band offset and valence band offset, respectively.

tical properties associated with broad applications [83–87]. For instance, Fig. 3.3 shows the transmission electron microscopy (TEM) images of Ge-core/Si-shell NWs synthesized by using a chemical-vapor deposition method applicable to a variety of nanoscale materials [83]. This structures shows a exceptionally high mobility of free carriers for the field-effect transistors.

Compared to the traditional planar bulk heterojunctions, core-shell NWs possess a few exceptional advantages. The radial-heterojunction structure contributes to a huge interface/volume ratio, making it extremely efficient for free-carrier transport and exciton dissociation in energy applications. Additionally core-shell NWs possess more degrees of freedom to tune their band offsets by varying structural geometry and corresponding quantum confinement and strain conditions. Moreover, the radial geometry makes the involved two semiconductors, i.e., the core and shell of the
wire inequivalent. Therefore, A-core/B-shell heterojunctions exhibit a different band lineup from those B-core/A-shell ones, suggesting that these NWs can be much more physically-rich than typical bulk heterojunctions.

Figure 3.3. Ge-core/Si-shell NWs. (a) Bright-field image of an Ge-core/Si-shell NW. Scale bar is 50 nm. (b), (c) Scanning TEM elemental maps of Ge (red) and Si (blue) concentrations, respectively. (d) High-resolution TEM image of a representative NW. Scale bar is 5 nm. (e) Elemental mapping cross-section showing the Ge (red circles) and Si (blue circles) concentrations. (f) High-resolution TEM image of an annealed Ge-core/Si-shell NW. Scale bar is 5 nm. (g) Elemental mapping cross-section gives a 5-nm shell thickness with a sharp interface consistent with the TEM image. [83]

All of these characteristics make core-shell NWs extremely attractive for nanoelectronics and renewable energy interests because excited e-h pairs can be efficiently separated under optimized band lineups. The separated free carriers can be extracted efficiently due to their observed large mobilities. Therefore, understanding the elec-
tronic structure and, more importantly, how to control band lineups are of central interest for researchers working on core-shell nanostructures.

To date there have been studies focusing on band gaps and band offsets of core-shell NWs [88–93], but with a limited quantitative understanding of how to control the band lineups, which has substantially impeded the further experimental progresses and potential applications. On the other hand, it is known that the electronic structure of NWs is sensitive to the strain condition [94–96]. Particularly for core-shell NWs, strain is intrinsic because of the lattice mismatch between core and shell lattices. More importantly, one exceptional advantage of NWs is that their small size renders them able to tolerate a relatively large lattice mismatch and corresponding strain without introducing significant defects to degrade the electrical performance of NWs [97,98]. These facts make strain ideal for controlling the band alignments of core-shell NWs.

In addition to strain, quantum confinement effects are known to significantly influence the electronic structure of NWs [99]. Therefore, it may be possible to vary the size and degree of quantum confinement of core-shell NWs in order to control the band gap and band alignments accordingly.

Motivated by these considerations, we perform first-principles calculations and investigate band offsets in Si/Ge core-shell NWs by applying the axial strain and varying the degree of quantum confinement, respectively. Our study shows that strain can substantially tune electronic band alignments of Si/Ge core-shell NWs. The applied axial strain has two impacts on the electronic structure of studied core-shell
NWs: first, it shifts the band energy of silicon and germanium states differently so that the band offset is varied significantly; secondly, it changes the charge distribution of silicon and germanium states, e.g., an original germanium-shell state can evolve into a silicon-core state under the appropriate strain, then band offsets are changed consequently. Furthermore, varying the size of Si/Ge core-shell NWs is efficient to modify both valence and conduction band offsets simultaneously. Our first-principles results provide useful information for optimizing band alignments of Si/Ge core-shell NWs.

3.2 First-Principles Simulation Setup

![Top View and Side View of Atomistic Structure](image)

Figure 3.4. The top view and side view of the ball-stick model of atomistic structure of a Si/Ge Core-shell NW along the (110) direction. The yellow balls represent silicon atoms, the green ones represent germanium atoms, and the white ones represent the surface passivated hydrogen atoms.

The calculation was performed within the framework of the density functional theory (DFT) using the local density approximation (LDA) [4, 7]. We employ norm-conserving pseudopotentials [12] with a plane-wave basis set. The energy cutoff for the
plane waves is 24 Ryd and the reciprocal space is sampled by a $1 \times 1 \times 16$ k-point grid. It has to be pointed out that DFT/LDA calculations usually cannot provide accurate band energy and gaps because many-body effects are not adequately included [29]. However, our predicted trends of band offsets resulting from the strain and quantum confinement are shown to be reliable.

In this work, we focus on Si/Ge core-shell NWs along the (110) direction, since SiNWs with a diameter less than 10 nm were found to be primarily along (110) direction in experiments [100]. Fig. 3.4 presents the ball-stick model of our calculated core-shell NWs. All surface dangling bonds are passivated by hydrogen atoms. The supercell method is applied with an approximate vacuum distance of 8 Å between neighboring NWs. All studied NWs are fully relaxed according to the forces and stress calculated in DFT/LDA. The calculation is conducted with the QUANTUM ESPRESSO package [79].

3.3 Electronic Structures in Si/Ge Core-Shell NWs

Presented in Fig. 3.5 is a typical band structure and charge distributions of Si/Ge core-shell NWs. Here we have to define band offsets in an operational way, following previous works [91]. For example, in Fig. 3.5 (b) the top of the valence band (TVB) (V1) and bottom of the conduction band (BCB) (C1) are apparently confined within the core and shell regions, respectively. Then we scan electronic states away from the band gap and calculate the corresponding projection of density of states (PDOS) for germanium and silicon atoms, respectively; if more than 75% of PDOS belongs to the
Figure 3.5. Band structure, band offset, and charge distributions of a typical Si-core/Ge-shell NW (with core diameter of 1.7 nm and shell diameter of 2.5 nm). (a) Band structure. (b) Top views of charge distributions of four electronic states marked by red point and blue points in (a) and the dashed circles show the core and shell sizes. (c) A schematics of the band alignment of this core-shell NW.
core atoms, we define this electronic state to be a confined core state; if more than 70% of PDOS belongs to shell atoms, we define this electronic state as a confined shell state; if the electronic state cannot be assigned to either core or shell state, we define it as a non-confined state, e.g., V3 (the third valence band) and C11 (the eleventh conduction band) in Fig. 3.5 (b). Then the band offset is defined by the energy difference between BCB or TVB state and the nearest non-confined state, which are marked by $\Delta E_c$ and $\Delta E_v$ as shown in Fig. 3.5 (a).

We choose the criterion for core states (75%) to be slightly higher than that of shell states (70%), because the ground-state wave function usually tends to distribute near center of NWs due to the cylindrical shape of NWs. It has to be pointed out that the value of band offsets depends on our established criterion, but the fundamental trend of the evolution of the band alignment is reliable.

For all core-shell NWs calculated in this work, we find the TVB is always a germanium state and the BCB is always a silicon state, which gives a type II band alignment as schematically shown in Fig. 3.5 (c). This is consistent with the type II band alignment observed in bulk Si/Ge heterojunctions [101].

3.4 Strain Control of Band Offsets in Si/Ge Core-Shell NWs

We first study how the applied axial strain modifies the band alignment of Si/Ge core-shell NWs. In Fig. 3.6, we have presented the variation of band offsets in four typical Si/Ge core-shell NWs under different strain conditions. The applied strain varies from -3% to 5%, which is far below the critical fracture for Si/Ge core-shell
NWs [102] and the critical strain needed for theoretically predicted phase transitions [96].

Figure 3.6. Band offsets of four strained Si/Ge core-shell NWs. (a) A Si-core/Ge-shell NW with core diameter of 1.2 nm and shell diameter of 2.5 nm. (b) A Si-core/Ge-shell NW with core diameter of 1.7 nm and shell diameter of 2.5 nm. (c) A Ge-core/Si-shell NW with core diameter of 1.3 nm and shell diameter of 2.5 nm. (d) A Ge-core/Si-shell NW with core diameter of 1.8 nm and shell diameter of 2.5 nm.

The prominent feature of Fig. 3.6 is that valence band offsets vary monotonically and significantly with the applied axial strain; the tensile strain always reduces the valence band offsets whose tunable range is about a few hundred meV. In particular, in Fig. 3.6 (b) and (c) the valence band offsets are even diminished for a tensile strain larger than 4%. This strain-sensitive effect of valence band offsets gives hope to tension sensors based on core-shell NWs because the corresponding electrical current crossing the radial heterojunction can be turned on or off by the strain and associated
band offsets. A tunable band offset is also preferred for photovoltaic applications because the efficiency of the separation of e-h pairs around the interface of core-shell NWs can be modified by the strain-controlled band offsets.

On the other hand, the conduction band offsets shown in Fig. 3.6 are not so sensitive to the applied axial strain. Their variations as induced by strain are not monotonic, and the fluctuations of the conduction band offsets are less than 0.2 eV although the absolute values are sensitive to the core and shell sizes and chemical components. Therefore, this result indicates that applying axial strain may not be an efficient choice to control the electron transfer in narrow Si/Ge core-shell NWs.

There are at least two mechanisms contributing to the above changes of band offsets of strained Si/Ge core-shell NWs. The first is the different response of the band energy of electronic states to the applied strain. As an example, we track the shift of the band energy in a Si-core/Ge-shell NW. We pick the TVB state (which is a confined shell state in the shell) and the first non-confined valence state. Their top views of the charge distributions are shown in Fig. 3.7 (a) and (b), respectively. The evolution of the energy difference between these two states with the applied strain is shown in Fig. 3.7 (c). The wavefunctions of these two states are found not sensitive to the external strain. However, it is clear that their energy difference is significantly reduced by the applied tensile strain. As a result, the corresponding band offsets are substantially changed. Interestingly, when the applied tensile strain is greater than 4%, the first non-confined state shown in Fig. 3.7 (b) surpasses the original TVB
state (Fig. 3.7 (a)) and becomes the highest valence state, making the valence band offset zero accordingly.

Figure 3.7. Band energy evolution with the applied axial strain of a Si-core/Ge-shell NW with a core diameter of 1.7 nm and a shell diameter of 2.5 nm. (a) and (b) Charge distributions of the TVB and the first non-confined valence state, respectively. (c) The energy difference between the plotted electronic state in (a) and (b) under strain.

Beyond confined and non-confined states, the different response to the applied strain is also observed between confined core and confined shell states, which means that this strain mechanism does not depend on the detailed definition of band offsets. According to our calculations, the different responses of electronic states to the applied strain explains the zero valence band offsets discovered in Fig. 3.6 (b) and (c), in which the valence band offsets gradually decrease to zero with an increasing applied tensile strain.
Figure 3.8. (Top views of charge distribution evolution with strains. This electronic state is picked from a conduction state near the band gap in a Si-core/Ge-shell NW (with core diameter 1.71 nm and shell diameter 2.47 nm). (a) 0% strain. (b) -3% strain. (c) 3% strain. (d) 5% strain.

The other mechanism contributing to the change of band offsets under applied strain is the variation of electronic wavefunctions themselves. As shown in Fig. 3.8, we have plotted the charge distribution of a conduction state near the band gap in a Si-core/Ge-shell NW. Without applied strain, this electronic state is a well-defined germanium shell state as shown in Fig. 3.8 (a). However, when the strain is applied and gradually increases, the spatial distribution of this electronic state is substantially modified. In particular, as shown in Fig. 3.8 (d), when the applied strain is around 5%, a significant amount of the charge moves to the core region of the NW. As a result, this state evolves into a non-confined state from an original shell state, which consequently changes the value of band offsets because of our definition. According
to our calculations, this variation of the charge distribution is the reason for those zero conduction band offsets discovered in Fig. 3.6 (d).

3.5 Geometry Control of Band Offsets in Si/Ge Core-Shell NWs

In addition to strain effects, the geometry of core-shell NWs can also be used to control band offsets. Since there are so many ways to vary the geometry of NWs, we focus on the simplest case in this study; the core size is fixed while the shell size is changed. The main results regarding the band offset are summarized in Fig. 3.9. First, Si-core/Ge-shell and Ge-core/Si-shell NWs display qualitatively different evolutions with a change of shell thickness. For Si-core/Ge-shell NWs, the conduction band offset decreases and the valence band offset increases with the growth of the whole wire size; for Ge-core/Si-shell NWs, the conduction band offset increases and the valence band offset decreases with the growth of the whole wire size. The reason of these trends is that the quantum confinement in the shell region is reduced if the whole wire size increases. Taking Si-core/Ge-shell NWs as an example, the energy gap of Ge shell states decreases since the confinement in the shell is reduced for a larger-sized shell, while the core-state gap does not change much because the core size is fixed. According to the schematics shown in Fig. 3.5 (c), the conduction Ge state (the upper red line) moves downward and the valence Ge state (the upper blue line) moves upward, while the conduction and valence Si states (the lower red line and the lower blue line) are nearly fixed. Thus, the conduction band offset decreases and the valence band offset increases.
3.6 Summary of this Chapter

In conclusion, band alignments in Si/Ge core-shell NWs are significantly modified by the strain condition and quantum confinement. Through our first-principles calculations, we specify that applying the axial strain is effective for tuning valence band offsets without significantly changing conduction band offsets. We also discuss the physical origins for these variations of band alignments in Si/Ge core-shell NWs. Moreover, the geometry variation provides another degree of freedom to effectively tune both conduction band offset and valence band offset by changing the quantum confinement of the core-shell nanostructures. Our proposed band offset engineering may be of practical importance for nanoelectronics and renewable energy research.
4. LATTICE VIBRATIONS IN SI/GE CORE-SHELL NANOWIRES

4.1 Introduction

In solids, the mass of a nucleus is several thousand times larger than that of an electron, resulting in a much smaller relaxation time for electrons compared to that of nuclei. Therefore, when studying the motion of the nuclei, we usually assume that electrons are in their equilibrium states for a given nuclear configuration. This is called the adiabatic approximation or the Born-Oppenheimer approximation [103]. Another consequence of the mass difference between electrons and nuclei is that wavefunctions of the nuclear component are spatially more localized than those of the electronic component, thus we can treat the motion of nuclei classically.

In crystal lattices, the displacement of one nucleus from its equilibrium position results in a recoil force acting on the nucleus, leading to a vibration around the equilibrium. When the displacement is small, the potential that the nucleus feels can be approximately treated as a harmonic potential and the corresponding vibration is the harmonic oscillation. Since the atoms are coupled in the solids, the vibration
of one atom will induce vibrations of other atoms nearby and the vibrations further spread to the whole crystal back and forth, resulting in collective vibrations of all atoms, as shown in Fig. 4.1. The quantum-mechanical description of these atoms oscillating at a single frequency is phonon.

Phonons are the fundamental description of lattice vibrations in solids. To a large extent, they define many finite-temperature properties of solids. For example, they play an important role for quite a number of microscopic and bulk phenomena, such as neutron scattering, inelastic nuclear absorption, infrared absorption, raman scattering, thermal expansion, and phase transitions. A deep understanding of the phonon behavior, especially that of newly fabricated materials, is a necessary condition for future applications.
As we have discussed in the previous chapter, core-shell nanowires (NWs), a type of novel quasi-one-dimensional (1D) nanostructures, have ignited tremendous interest to date because these unique radial heterojunctions provide extra degrees of freedom to tune their mechanical, electric and optical properties, making them superior to usual homogenous materials [83,97,104,105]. Among numerous core-shell structures, Si/Ge core-shell NWs are of broad interest because of the known importance of silicon in microelectronics. In particular, recent research has shown that Si/Ge core-shell NWs possess a number of unusual thermal transport features for energy applications [106–108]. For example, the enhanced differences in thermal and electric transport behaviors in the core and shell make core-shell NWs promising for higher-efficiency thermoelectric devices [109].

In order to realize and optimize the above interesting properties, it is imperative to determine how to obtain structural information of NWs and, thereafter, how to design core and shell geometries. Traditional approaches, such as the transmission electron microscopy (TEM) and neutron scattering measurements, are time consuming or expensive for the large-scale production. Thus alternative convenient ways are highly desirable. On the other hand, lattice vibrational modes are known to be sensitive to the chemical bonding and boundary conditions of nanostructures and they are readily accessible by various experiments, such as Raman scattering [110–112], making it possible to study lattice vibrational modes of core-shell NWs and extract useful structural information conveniently.
Beyond application interests, lattice vibrational modes and corresponding phonons are fundamental properties of solids. However, compared with bulk semiconductors, there have been very limited first-principles studies on vibrational modes of core-shell NWs. How quantum confinement and internal strain in core-shell NWs affect their phonons and thermal properties has not been understood well yet. Therefore, calculating lattice vibrational modes in Si/Ge core-shell NWs will be an appropriate starting point for understanding electric and thermal properties of general reduced-dimensional structures.

In this work, we employ first-principles calculations and focus on the following two topics of Si/Ge core-shell NWs: 1) how the unique core-shell structure modifies lattice vibrational modes and what their impacts on thermal properties are; 2) how to extract the structural information from the characteristics of lattice vibrational modes and provide clues for experiments to design this type of core-shell materials.

Through our simulations, we find that the structural information of core-shell NWs can be efficiently obtained by frequency shifts of high-energy optical vibrational modes. This type of frequency shift can be explained well by the change of the strain induced by the lattice mismatch between the core and shell. Moreover, we have calculated the Raman scattering spectrum of Si/Ge core-shell NWs, showing that the above frequency shifts can be detected by experiments. Meanwhile, we have identified the existence of radial breathing modes (RBMs) in Si-core/Ge-shell NWs, whose frequency variation according to the size change of NWs can be understood by an elastic media model.
The rest of this chapter is organized as the following. In Sec. 2, we present our calculating methodology and computational setup. In Sec. 3, we discuss the simulation results, including properties of high-frequency optical modes, transverse and longitudinal optical modes splitting, Raman scattering spectra, and the RBM frequency dependence on the structure of core-shell NWs. In Sec. 4, we summarize our studies and conclusion.

4.2 First-Principles Simulation Setup

Our calculation is based on the density functional theory (DFT) within the local density approximation (LDA) [4,7] as implemented in the Quantum ESPRESSO package [79]. We employ the plane-wave basis and the pseudopotential approximation [12]. The plane-wave energy cutoff is set to be 16 Ryd. For the Brillouin zone (BZ) integration, we use a $1 \times 1 \times 8$ k-point sampling grid. All calculations are done in a supercell arrangement [113] with a 0.7-nm vacuum distance between neighboring NWs. Our studied NWs are oriented along the (110) direction, which is an energetically preferred direction [100]. Dangling bonds on the surface of NWs are passivated by hydrogen atoms. All NWs are fully relaxed according to the force and stress calculated under DFT/LDA.

There are a number of degrees of freedom to decide the structure of core-shell NWs, such as the core size, shell size and the chemical compositions. Because it is impossible to study all of them in one project, here we investigate Si/Ge core-shell NWs with a fixed total diameter ($\sim 2.3$ nm) but a varying core-shell size ratio. At the
same time, homogenous NWs are also studied for comparison purposes. The detailed structural information of our studied NWs is reported in Table 4.1.

<table>
<thead>
<tr>
<th>NWs</th>
<th>Core</th>
<th>Shell</th>
<th>Core diameter (nm)</th>
<th>Shell diameter (nm)</th>
<th>Lattice constant (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si-1</td>
<td>Si</td>
<td>Si</td>
<td>2.24</td>
<td>2.24</td>
<td>0.378</td>
</tr>
<tr>
<td>GeSi-1</td>
<td>Ge</td>
<td>Si</td>
<td>1.12</td>
<td>2.21</td>
<td>0.383</td>
</tr>
<tr>
<td>GeSi-2</td>
<td>Ge</td>
<td>Si</td>
<td>1.75</td>
<td>2.30</td>
<td>0.390</td>
</tr>
<tr>
<td>Ge-1</td>
<td>Ge</td>
<td>Ge</td>
<td>2.41</td>
<td>2.41</td>
<td>0.401</td>
</tr>
<tr>
<td>SiGe-1</td>
<td>Si</td>
<td>Ge</td>
<td>1.09</td>
<td>2.33</td>
<td>0.396</td>
</tr>
<tr>
<td>SiGe-2</td>
<td>Si</td>
<td>Ge</td>
<td>1.66</td>
<td>2.29</td>
<td>0.388</td>
</tr>
<tr>
<td>Si-2</td>
<td>Si</td>
<td>Si</td>
<td>1.61</td>
<td>1.61</td>
<td>0.382</td>
</tr>
<tr>
<td>Ge-2</td>
<td>Ge</td>
<td>Ge</td>
<td>1.76</td>
<td>1.76</td>
<td>0.405</td>
</tr>
</tbody>
</table>

We employ the linear response approach to obtain lattice vibrational modes and their frequencies [51–54]. Only vibrational modes at the Γ point of the first BZ is calculated because it is of particular interest for the first-order Raman scattering measurement.

We calculate the Raman scattering spectrum by considering the non-resonant first-order process, which can be described within the Placzek approximation [114, 115]. Then the Raman activity $I_{\text{ramman}}^k$ associated with the vibrational mode $k$ is

$$I_{\text{ramman}}^k = |e_s \cdot \alpha^k \cdot e_L|,$$

where $e_s$ is the polarization of incoming photon and $e_L$ is the polarization of scattered outgoing photon, respectively.
The Raman tensor $\alpha^k$ is defined as

$$\alpha^k = \sqrt{\Omega} \sum_{al} \frac{\partial \chi_{ij}}{\partial r_{al}} u^k_{al},$$  \hspace{1cm} (4.2)$$

where $\chi_{ij}$ is the electric polarizability tensor, $a$ is the index to specify the atom in the unit cell, and $l$ is the index of coordinates. $u^k_{al}$ is the displacement of the atom $a$ along $l$ direction in the vibrational eigenmode $k$.

For realistic experimental cases, NWs are usually oriented randomly. Because of the depolarization effect, we may only consider the circumstances that the direction of the incident beam, the polarization direction of this beam, and the direction of the observation are perpendicular to each other. Then the spatially averaged Raman activity of those randomly orientated NWs is given by:

$$I_{\text{ramman}}^k = 5(\alpha^k_{xx} + \alpha^k_{yy} + \alpha^k_{zz})^2 + \frac{7}{4}[\alpha^k_{xx} - \alpha^k_{yy}]^2 + (\alpha^k_{xx} - \alpha^k_{zz})^2 + (\alpha^k_{yy} - \alpha^k_{zz})^2 + 6(|\alpha^k_{xy}|^2 + |\alpha^k_{xz}|^2 + |\alpha^k_{yz}|^2)].$$ \hspace{1cm} (4.3)$$

Finally, because our calculated Raman spectra only consider the first-order process, they may be substantially different from some Raman experiments which could be dominated by resonant effects. However, our obtained Raman activity shall still be useful to understand those resonant experiments, such as their selection rules.
4.3 Vibrational Modes in Si/Ge Core-Shell NWs

4.3.1 High-Frequency Optical Modes

The density of vibrational modes (DVM) at the Γ point is presented in Fig. 4.2. It has to be pointed out that those vibrational modes mainly consisting of hydrogen-atom motions are excluded from our plots because we are interested in intrinsic vibrational modes of NWs that are not sensitive to the environment. Therefore our results and conclusion shall be more universal and less affected by complicated passivations as observed in various experiments.

The most prominent feature of these DVMs is that there are two characteristic peaks among most core-shell NWs, which are marked by \( O_{Si} \) and \( O_{Ge} \) in Fig. 4.2, respectively. To understand the physical origin of these peaks, we have plotted the real-space vibration of these modes corresponding to peaks \( O_{Si} \) and \( O_{Ge} \) in Fig. 4.3.

The first character of these modes is that they are mainly confined within the core or shell, respectively. In another word, these Si and Ge vibrational modes are decoupled. For example, Fig. 4.3 (a) is mainly the Si-Si vibrational mode confined within the core regime while Fig. 4.3 (b) is mainly the Ge-Ge mode within the shell regime. This decoupling is due to the significant mass difference between Si and Ge atoms and the bonding strength differences between Si-Si and Ge-Ge bonds.

The second character is that the neighboring atoms of these modes plotted in Fig. 4.3 have opposite vibrational directions, a typical feature of optical modes of semiconductors at the Γ point. Based on the above two characters, we conclude that
Figure 4.2. Vibrational mode density of states of Si/Ge core-shell NWs and homogenous NWs with wire diameter of about 2.3 nm in Table 4.1. (a) Si-1 NW (b) GeSi-1 NW (c) GeSi-2 NW (d) Ge-1 NW (e) SiGe-1 NW (f) SiGe-2 NW. The red dashed line is used to guide the shift trend of the peak of high-frequency Ge-Ge optical modes.
these peaks $O_{Si}$ and $O_{Ge}$ originate from those high-frequency optical modes of bulk Si and Ge. This can be further confirmed by their frequencies; the $O_{Ge}$ peak has a frequency around 8 THz while the $O_{Si}$ peak has a frequency around 15 THz. They are consistent with frequencies of optical modes in Si and Ge bulk crystals.

However, the frequencies of peaks $O_{Ge}$ and $O_{Si}$ are not exactly the same as their bulk counterparts. Instead they exhibit substantial frequency shifts as the core-shell size varies as shown in Fig. 4.2. For example, when we expand the Ge-core size from zero to the whole Ge-core/Si-shell wire (Figs. 4.2 (a) to (d)), both peaks $O_{Si}$ and $O_{Ge}$ have a red shift. Similarly, both peaks have a blue shift as Si core size increases in the case of Si-core/Ge-shell NWs, as shown in Figs. 4.2 (d) to (f).

The physical origin of the above shifts of prominent peaks in the DVM may be attributed to two factors, quantum confinement and strain condition. First, quantum confinement in such narrow NWs definitely modifies the frequency of lattice vibra-
tional modes as discovered by previous studies [110–112]. However, since the whole diameter of all studied NWs is the same, quantum confinement shall not be a major factor for the frequency shift observed in Fig. 4.2. On the other hand, the internal strain of core-shell NWs induced by lattice mismatch can be an important cause for those frequency shifts. In our Si/Ge core-shell NWs, the Si part is stretched while the Ge part is compressed. As the core-shell size changes, the average lattice constant along the axial direction varies according to the Vegard’s law \(a = a_A^0(1 - X) + a_B^0X\), where \(a\) is the lattice parameter of \(A_{1-X}B_X\) crystal, \(X\) is the concentration of constituent element B, \(a_A^0\) and \(a_B^0\) are the lattice parameters of pure A and pure B crystals, respectively.), which can be seen from Table 4.1. As a result, the intrinsic strain will gradually change as we vary the core-shell size ratio, resulting in significant frequency shifts of the Si and Ge modes.

To further confirm the above explanation, we focus on the frequency shifts of particular optical modes picked from \(O_{Si}\) and \(O_{Ge}\) peaks, respectively. As shown in Fig. 4.4, for both Si-core/Ge-shell and Ge-core/Si-shell NWs, these frequency shifts are approximately linear in the square of the ratio of the core and NW radii, which is proportional to the ratio of the number of atoms in the core and the whole NW. Since this linear relation is exactly the lattice constant variation according to the Vegard’s Law, the change of the frequency shall be a result of the variation of the internal strain.

Meanwhile, we calculate homogenous Si NWs (SiNWs) and Ge NWs (GeNWs) with applied axial strain to check the strain effect as shown in Fig. 4.5. The size
Figure 4.4. The highest frequency of the optical mode picked from $O_{Si}$ and $O_{Ge}$ peaks. (a) Si-core/Ge-shell NWs with shell diameter of about 2.3 nm and core size varying. (b) Ge-core/Si-shell NWs with shell diameter of about 2.3 nm and core size varying.

Figure 4.5. Highest optical mode frequencies with external uniaxial strains applied to pure Si NW and pure Ge NW respectively. (a) Si-2 NW, the cross marks the frequency position of the highest Si optical mode in SiGe-2 NW. (b) Ge-2 NW, the cross marks the frequency position of the highest Ge optical mode in GeSi-2 NW.
of homogenous NWs is chosen to be similar to the corresponding core size of Si/Ge core-shell NWs (SiGe-2 and GeSi-2 NWs listed in Table 4.1). Here we focus on the highest-frequency optical mode, which is the same mode plotted in Fig. 4.4. For both SiNWs and GeNWs, these modes exhibit a monotonic change with the applied axial strain; their frequencies increase with the compressive strain and decrease with the tensile strain, which is consistent with the trend of the peak shifts in core-shell NWs. Moreover, we mark the highest-frequency optical mode of the corresponding core-shell NWs in Fig. 4.5. We find that their frequencies are close to those of the highest-frequency optical mode of homogenous NWs under the similar strain. Therefore, the variation of the strain condition is the main reason for the frequency shifts of high-frequency optical modes in core-shell NWs. This result builds a bridge connecting the frequency shift of characteristic optical modes with the structure of core-shell NWs.

4.3.2 Transverse and Longitudinal Optical Modes Splitting

The strain effect not only causes frequency shifts of optical modes but also affects the transverse and longitudinal optical modes splitting (TO-LO splitting). This TO-LO splitting is observed in Si/Ge core-shell NWs, as shown in Fig. 4.6. Here we only plot the highest-frequency TO, which is always vibrating along the (001) direction. In Fig. 4.6, the splitting varies as the core shell structure changes. This change in splitting is a result of competition of internal strain and quantum confinement.

In previous studies [112], the quantum confinement effect will enhance the TO-LO splitting. As a result, we observe this splitting for homogenous NWs in Fig. 4.6,
where \( r/R = 1 \). On the other hand, for example, in Fig. 4.6 (a), as the size of the Si core increases, the strain has different impacts on TO and LO. As a result, the TO-LO splitting is no longer a simple monotonic relation to the core-shell geometry.

Figure 4.6. Highest TO-LO mode splitting in core-shell NWs. (a) core modes splitting in Si-core/Ge-shell NWs. (b) core modes splitting in Ge-core/Si-shell NWs.

4.3.3 Raman Scattering Spectrum

Raman scattering is a widely used approach for detecting lattice vibrational modes of solids. If the above frequency shifts of characteristic optical modes can be observed in their Raman scattering spectrum, the detection of these modes will be useful to understand the structures of core-shell NWs. Following the formulas in Eq. (4.2) and Eq. (4.3), we calculate the first-order Raman scattering spectra of our studied Si/Ge core-shell NWs and present them in Fig. 4.7.
In Fig. 4.7, we can see the following important features. First, those highest-frequency optical modes in the peaks $O_{Si}$ and $O_{Ge}$ are strongly active in the Raman scattering spectra. This is reasonable because they originate from bulk optical modes that are known to be highly Raman active. Since we have shown that these shifts are caused by the strain and associated with structural variations, the active Raman scattering peaks provide a convenient way to obtain structural information and the corresponding strain condition inside Si/Ge core-shell NWs.
We also observe that the Raman scattering signal intensity strongly depends on the type of atoms involved in the vibrational modes. As shown in Fig. 4.7, those Ge modes have much stronger Raman signals than those Si modes. This can be attributed to the larger size of Ge atoms and higher charge mobility, so that their motions can induce much stronger changes of polarizability, leading to enhanced Raman scatter signals.

Meanwhile, the intensity of Raman scattering signals also depends on the spatial location of those atoms involved in the vibrational modes. The Raman scattering peak is usually enhanced if the corresponding mode is within the core region while it is depressed if the mode is within the shell region. For example, though SiGe-2 (Si core) NW has a lower proportion of Si atoms than GeSi-1 (Si shell) NW as listed in Table 4.1, the relative Raman signal of the highest-frequency Si mode in SiGe-2 NW is stronger than that in GeSi-1 NW, as shown in Figs. 4.7 (b) and (d). A similar result can be observed in other NWs if we compare Figs. 4.7 (a) and (e).

Other than core-shell NWs, if comparing the Raman spectra of the homogenous SiNW and GeNW with the similar size as shown in Figs. 4.7 (c) and (f), we find that the SiNW, different from the GeNW which has only one dominant Raman peak at highest optical frequency, displays bright Raman activities for many phonon modes with lower frequencies. Since both bulk Si and Ge have only one dominant peak in the first-order Raman scatter spectrum, those newly active low-frequency modes in SiNWs imply that these vibrational modes are more affected by the quantum confinement and symmetry breaking in SiNWs. This has been observed by recent
experiments [110, 111]. Such a different variation of vibrational modes may be helpful to explain why the thermal conductivity of SiNWs differs from their bulk crystals more than that of GeNWs. [108].

4.3.4 Radial Breathing Modes

After discussing the high-frequency optical vibrational modes in Si/Ge core-shell NWs, we turn to another type of collective mode: the radial breathing modes (RBM). Because of the unique geometry of quasi-one-dimensional nanostructures, their RBM are of great interests and importance. For example, they are characteristic modes reflecting the geometry of carbon nanotubes [116, 117]. Recent research [112, 119] has also shown that the RBM are Raman active in narrow semiconducting NWs and they are closely related to the diameter of NWs. Then an obvious question is whether we can observe similar RBM in core-shell NWs.

Our first-principles result shows that the answer for core-shell NWs is more complicated than usual homogenous nanostructure cases. For example, the RBM are identified in the Si-core/Ge-shell NWs studied in this work as shown in Fig. 4.8 (b). Moreover, we find that these RBM are Raman active and can be detected by Raman scattering experiments, as indicated in Fig. 4.7 (a), (b) and (c). On the other hand, for the other type of NWs, Ge-core/Si-shell NWs, we cannot identify the RBM simply by sight. In fact, even for GeNWs with a similar diameter as the core-shell NWs we studied, no RBM can be easily identified. The relative softer bonding between Ge atoms and the relatively larger mass of Ge atoms substantially reduce the frequencies
of the RBM. As a result, the RBM will mix with other low-frequency modes because of the application of sum rule in our simulations.

The most important motivation for studying the RBM is to see if they have a relation with the structure of nanomaterials, such as the diameter of NWs. Previous studies [112] have successfully applied the continuous elastic media model to study the RBM in SiNWs, encouraging us to apply a similar approach to study the RBM in Si-core/Ge-shell NWs. By using the classic wave equation with a cylindrical boundary condition, and regarding core and shell regions as two different homogenous elastic media with different sound velocities $c_1$ and $c_2$ respectively, the vibration of RBM can be described as $\vec{u}(x, y, z) = \vec{u}(x, y) = u(r)\hat{r}$, which satisfies the following equations:
\[
\frac{\partial^2 u(r)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r)}{\partial r} + \left( \frac{\omega^2}{c_1^2} - \frac{1}{r^2} \right) u(r) = 0 \quad \text{for} \quad 0 \leq r \leq R_1, \quad (4.4a)
\]
\[
\frac{\partial^2 u(r)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r)}{\partial r} + \left( \frac{\omega^2}{c_2^2} - \frac{1}{r^2} \right) u(r) = 0 \quad \text{for} \quad R_1 \leq r \leq R_2, \quad (4.4b)
\]

where \( \omega \) is the frequency of RBM. \( R_1 \) is the core radius and \( R_2 \) is the shell radius, \( u_1(r) \) and \( u_2(r) \) are RBM wavefunctions in the core and shell region, respectively.

We solve the equations, Eqs. (4.4), using the free boundary condition at \( r = R_2 \) and the continuous and differentiable continuous condition at \( r = R_1 \), we get an equation Eq. (4.5) which determines \( \omega \), the frequency of RBM:

\[
c_1 J_1(\frac{\omega R_1}{c_1}) \left[ Y_1'(\frac{\omega R_1}{c_2}) J_1'(\frac{\omega R_2}{c_2}) - J_1'(\frac{\omega R_1}{c_2}) Y_1'(\frac{\omega R_2}{c_2}) \right] \\
c_2 J_1'(\frac{\omega R_1}{c_1}) \left[ Y_1(\frac{\omega R_1}{c_2}) J_1(\frac{\omega R_2}{c_2}) - Y_1'(\frac{\omega R_1}{c_2}) J_1'(\frac{\omega R_2}{c_2}) \right] = 0 \quad (4.5)
\]

where \( J_1 \) and \( Y_1 \) are first-order Bessel functions of the first kind and second kind, respectively. By applying the perturbation theory and the approximation of Bessel’s functions in Eq. (4.6) when \( x \gg \frac{3}{4} \) (this is satisfied in our simulations), we get

\[
J_1(x) \approx \sqrt{\frac{2}{\pi x}} \cos(x - \frac{3}{4} \pi), \quad (4.6a)
\]
\[
Y_1(x) \approx \sqrt{\frac{2}{\pi x}} \sin(x - \frac{3}{4} \pi). \quad (4.6b)
\]

A relation about the frequency of the RBM in core-shell NWs is obtained:
Therefore, the frequency $\omega$ of the RBM is determined by the ratio of core/wire size and that of the sound velocities of core and shell. The constant in Eq. (4.7) are a series of values of $x$ when $J'_1(x) = 0$, according to the free boundary condition. From the equation Eq. (4.7), we can easily reduce it to the known RBM frequency of pure NWs:

$$\omega \left( \frac{R_2}{c_2} + \frac{R_1}{c_1} - \frac{R_1}{c_2} \right) \approx constant$$  \hspace{1cm} (4.7)

We have applied the equation Eq. (4.7) to fit the data of RBM in the Si-core/Ge-shell NWs we study, as shown in Fig. 4.8(a). From the fitting curve, the phonon group velocity ratio in Si core and Ge shell is about 1.82, which is larger than the ratio in homogenous NWs or bulk materials ($\sim 1.56$). The enhancement of the phonon group velocity ratio in core-shell NWs is a result of intrinsic strain effect from the lattice mismatch of Si and Ge. For example, in Si/Ge core-shell NWs, the Si component is stretched while the Ge component is compressed, which further enhances the difference between the thermal transport behaviors of core and shell. This may be of interest for thermoelectric applications of core-shell NWs.
4.4 Summary of this Chapter

In summary, we have studied lattice vibrational modes of Si/Ge core-shell NWs by a first-principles approach. The frequency of high-energy Si-Si and Ge-Ge optical modes exhibits significant shifting when the core-shell size varies. Our analysis shows that these shifts are related to the variation of the strain conditions inside core-shell NWs. At the same time, the RBM is identified in the Si-core/Ge-shell NWs, whose frequencies strongly depend on the geometry of NWs, and our elastic media model explains well the frequency dependence of the RBM. Moreover, we have performed the calculation to obtain the first-order Raman scattering spectra of relevant NWs, in which the shift of high-frequency optical modes and the RBM can be identified, providing a convenient way to detect the structural information of core-shell NWs. Our studied vibrational modes and their frequencies could be useful for thermoelectric applications.
5. EXCITON SPECTRA OF TWO-DIMENSIONAL SEMICONDUCTING CARBON STRUCTURES

5.1 Introduction to Excitons and Exciton Spectra

An exciton is an important type of electronic excited state that influences the electric and optical properties of semiconductors. As shown in Fig. 5.1, an incident photon can excite a negatively-charged electron to the conduction band of semiconductors, leaving a positively-charged hole in the valence band. Then the attractive coulomb interaction brings these excited particles into an electron-hole (e-h) pair, forming an exciton. For these bound states, the sequence of binding energies is defined as the exciton spectrum, which is similar to the Rydberg energy sequence of the electronic states observed in hydrogen atoms.

The exciton spectrum is of particular importance for fundamental studies of excitonic effects. For example, the essential e-h pair concept and the observed $\frac{1}{n^2}$ Rydberg relation of the energy sequence ($n$ is the index of exciton states) motivates the hydrogenic model of excitons, which has been widely used to interpret experimental...
spectroscopy results. Moreover, the degeneracy and energy spacing between those bound states reflect the screened Coulomb interaction and the symmetry variations.

Secondly, exciton spectra of low-dimensional semiconductors are of particular importance in light of their unique anisotropic quantum confinement, which not only affects the optical spectrum but also modify the e-h interactions and corresponding binding energies. However, although e-h interactions are known to be dramatically enhanced in reduced dimensional structures [120–125], very few experiments can quantitatively measure this enlarged e-h binding energy in nanostructures. In fact, how these unique quantum confinements influence the exciton spectra and how one subsequently modifies the corresponding exciton models are largely unknown yet. Experimentalists can easily obtain the optical spectrum but it is extremely challenging to obtain the quasiparticle band gap at the same time, making it very hard to directly derive the binding energy of e-h pairs. As a result, based on the knowledge of exciton spectra, numerous exciton models for one-dimensional (1D) nanostructures [126–128]
and quantum wells [129–133] have been proposed, which give hope to explaining experimental results without needing costly simulations. However, there have been no such effort on 2D semiconductors.

Recently many-electron effects and optical properties of graphene and its derivatives have ignited substantial interests because of their enhanced many-electron effects [134–137]. However, other than studies of the optical absorption, the exciton spectra of these novel materials are not touched at all. Therefore, we are unable to extract the general features of e-h interactions and build appropriate exciton models of these confined two-dimensional (2D) systems.

Meanwhile, the first-principles simulation based on the many-body perturbation theory (MBPT) is particularly useful to solve the aforementioned problems because this reliable calculation methodology can provide the binding energy spectrum of excitons (including dark and bright states), their optical activities, and even their wavefunctions, at the quantum-mechanical level. This motivates us to employ MBPT to calculate excitonic effects of important derivatives of graphene, i.e., graphyne [138–143] and graphane [144–148]. First, we will reveal the unknown exciton spectra of these novel 2D structures; secondly, we will build a model that can quantify excitonic effects of more general 2D semiconductors without costly ab initio simulations, e.g., extrapolating the e-h binding energy, which is hard to measure directly in experiments.

Beyond fundamental scientific motivations, graphyne, a novel allotrope of graphene, is of particular interest for optical applications. Unlike other graphene derivatives,
such as graphane and fluorographene, whose low-energy optical transitions are depressed by the tetrahedral symmetry [134], the low-energy optical activity of graphyne may be prominent because of its planar atomistic structure and corresponding active inter-band transitions between $\pi$ electronic states [149]. Large-scale graphyne has not been fabricated to date despite substantial synthesis advances [149–155]. In this stage, a quantitative prediction of electric and optical properties of graphyne is crucial to foresee potential applications and motivate more research efforts.

In this chapter, we begin by revealing excited-state properties of a graphyne structure which is of current interest for fabrication. The quasiparticle (QP) band gap is appreciable (1.4 eV); the lowest-energy optical absorption peak is located at 1.0 eV, meaning a 400-meV $e-h$ binding energy; the near-infrared optical absorbance is more than 6%, making our studied graphyne one of the most efficient optical absorbers among known materials; this graphyne structure possesses an impressive spin singlet-triplet splitting ($\sim 150$ meV) of excitons. These features promise exciting energy and biomedical applications.

Moreover, based on our calculated exciton spectrum, we propose a modified one-parameter hydrogenic model in which the Coulomb potential is revised to capture the anisotropic quantum confinement of such a 2D semiconductor. To justify this model, we have applied it to graphyne and graphane, achieving excitonic spectra consistent with \textit{ab initio} results. Therefore, this model may provide a convenient way to estimate the exciton binding energy without knowledge of the QP band gap, which shall be of
broad interest to identify many-electron effects from the optical spectroscopy of 2D nanostructures.

5.2 First-Principles Simulation Setup

We employ density functional theory (DFT) and subsequent many-body perturbation theory (MBPT) to reveal electronic structures and optical properties of the graphyne and graphane structures. The ground-state calculations using local density approximation (LDA) are performed by the Quantum Espresso package [79] with a plane-wave basis and norm-conserving pseudopotentials. The plane-wave energy cutoff is set to be 60 Ry. A coarse $16 \times 16 \times 1$ k-point sampling of the first Brillouin zone (BZ) is applied to compute the self-energy within the single-shot $G_0W_0$ approximation [29] with a layered Coulomb truncation. A fine k-grid $(64 \times 64 \times 1)$ is interpolated from the coarse grid $(16 \times 16 \times 1)$ to obtain the converged excitonic states and optical absorption spectrum by solving the Bethe-Salpeter Equation (BSE) [30]. The GW/BSE calculations are done with the general plasmon-pole model [163] implemented in the BerkeleyGW package. Four valence bands and four conduction bands are included to calculate optical absorption spectra of the incident light polarized parallel to the graphyne plane because of the depolarization effect [120,125].

Usually, the imaginary part of the dielectric function, $\varepsilon_2(\omega)$, is defined as the optical absorption of semiconductors. However, for reduced dimensional structures, $\varepsilon_2(\omega)$ is not well-defined because it is inherently a bulk property. Actually, this is revealed by the fact that $\varepsilon_2(\omega)$ obtained from the BerkeleyGW code depends on the
size of the unit cell which has a vast vacuum around. In order to eliminate this artificial effect, we calculate the polarizability of the material and transfer it into the absolute optical absorbance as

$$\alpha(\omega) = \varepsilon_2(\omega) \cdot \omega \cdot \frac{d}{c}, \quad (5.1)$$

where \(\varepsilon_2(\omega)\) is the imaginary part of macroscopic dielectric function calculated from BSE, \(\omega\) is optical frequency, \(d\) is supercell size perpendicular to the planar structure, and \(c\) is the speed of light. Here we assume the thickness of graphyne is negligible and the detailed derivation can be found in Ref. [125].

### 5.3 Electronic Structures and Optical Properties of Graphyne

![Graphyne Structure and Band Structures](image)

Figure 5.2. (a) Top view of the ball-stick model of our studied graphyne structure. (b) DFT and QP electronic band structures. The black dots represent the DFT result and the blue curves are the QP band structure. The top of valence band from both calculations is always set to be zero.
The graphyne structures is shown in Fig. 5.2 (a). It is still a triangular lattice but without the honeycomb structure. The benzene rings are connected by -C≡C- chains. Compared with graphene, the unique character of this material is that it is a direct-gap semiconductor [143], a signature for intriguing optoelectronic properties.

The DFT and QP band structures of graphyne are presented in Fig. 5.2 (b), respectively. Because of the depressed screening in such a 2D semiconductor, enhanced self-energy correction enlarges the band gap from the DFT predicted 0.43 eV to 1.4 eV. The similarly enhanced many-electron correction is also observed in other 2D semiconductors [124,149]. At the same time, the desired direct band gap is kept at the M point even after the GW correction.

The optical absorption spectra of graphyne are presented in Fig. 5.3 (a). In the single-particle absorption spectrum without e-h interactions included (the blue curve), the optical absorption edge starts from the QP band gap (∼1.4 eV) due to the direct-gap nature. More interestingly, a huge optical absorbance is observed. For example, within the near-infrared and visible frequency regime, more than 6% of the incident light will be absorbed by a single atomic layer, making our studied graphyne to be one of the most efficient optical absorbers. This huge optical absorbance is a result of the significant overlap between the valence and conduction π electronic states in such a confined structure and consequently enhanced dipole transitions [30].

After including e-h interactions, we observe dramatic excitonic effects on the optical absorption spectrum as shown in Fig. 5.3 (a) (the red curve). First, two new absorption peaks (A1 and A2) appear below the QP band gap because of the for-
Figure 5.3. (a) Optical absorption spectra of graphyne with and without e-h interaction included. The absorbance value is obtained according to Ref. [30]. A 0.05 eV Gaussian broadening is applied to obtain these optical absorption spectra. (b) Exciton spectra of bound excitons. The black lines represent dark states and those red lines represent bright excitons. (c) Transition strength of excitons in graphyne.
mation of e-h pairs (excitons). In particular, the most prominent exciton with the lowest energy is located at 1.0 eV, implying a 0.4-eV e-h binding energy, which is an order of magnitude larger than those of excitons in bulk semiconductors. These enhanced excitonic effects are due to the substantially depressed screening and quantum confinement [120,122,124].

Moreover, we have calculated the spin-triplet excitons that are usually dark in the single-photon optical absorption spectrum due to the selection rule. The lowest-energy spin-triplet exciton is located at 0.85 eV in the optical spectrum, which is 150 meV below the first bright singlet exciton (A1) that is located at 1 eV. Such an enhanced spin singlet-triplet splitting (∼ 150 meV) is around an order of magnitude larger than those of typical semiconductors and even carbon nanotubes [156]. Since the spin singlet-triplet splitting is decided by the e-h exchange interaction (see Eq. (2.91) in Chapter 2) [30], the tremendous splitting observed in graphyne results from the significant overlap of electron and hole wavefunctions, which is consistent with the aforementioned huge optical absorbance.

The above unique optical properties of graphyne may give hope to numerous important applications. For example, the bright exciton A1 located at 1.0 eV falls into the ideal absorption edge of photovoltaic (PV) materials according to the Shockley-Queisser criterion [157], meaning it can absorb most solar light while keep the high output voltage. Moreover, the significant e-h binding energy (∼ 400 meV) and impressive spin singlet-triplet splitting results in possible long lifetimes for the lowest-energy excitons. In particular, because spin-triplet excitons are forbidden by the dipole tran-
sition, they may possess a much longer lifetime than singlet excitons. This will reduce
the exciton recombination and extend their diffusion length, giving a promising high
PV efficiency [158,159].

Other than energy applications, our studied graphyne structure exhibits an ex-
tremely strong absorbance between 1 eV and 2 eV, which is of particular interest for
biomedical applications. For instance, most biology tissues are transparent to the in-
cident light with a frequency between 1.3 to 1.9 eV [160]. Therefore, graphyne-based
devices can work as efficient optical absorbers for photo-therapy and bio-imaging
purposes.

5.4 Exciton Spectra of Graphyne

Beyond focusing on these optically prominent excitons, we have to obtain the
whole exciton spectrum, which is crucial for subsequent modelling efforts. In Fig. 5.3
(b) we list all bound exciton states of graphyne according to their binding energies.
It has to be pointed out that because the direct band gap is located at the M point
of the first BZ and there are three equivalent M points in the hexagonal structure of
the first BZ, each exciton energy level shall be triply degenerate. However, we do not
find any interactions of between these degenerate exciton states; it is safe to consider
only one set of excitons originating from one M point.

The exciton spectrum listed in Fig. 5.3 (b) immediately raises several interesting
questions. For example, the second lowest-energy level is doubly degenerate and
it is dark in the optical absorption spectrum. These features substantially conflict
with the widely used hydrogenic exciton model in which the lowest two excitons are non-degenerate and bright $1s$ and $2s$ states, respectively.

In order to understand this unusual exciton spectrum, we focus on their real-space wavefunctions. In Fig. 5.4, the six lowest-energy excitonic states are plotted (for degenerate states, we only plot one of them) in real space. First, we see the fundamental characters of these exciton wavefunctions are similar to the hydrogenic model, e.g., the spherical symmetry of the $s$ orbital, those angular momentum characters of $p$, $d$ and $f$ orbitals, and their nodal structures. As a result, we can identify these states...
with the same parities as the hydrogenic model, \textit{i.e.}, 1s, 2s, and 2p, \textit{etc.}, as marked in Fig. 5.3 (b).

Secondly, the optical selection rules on these states are also almost preserved. For example the \textit{s} states are bright while the \textit{p} states are dark. The only exception is the \textit{3d} states, which is dark while they are slightly bright in Fig. 5.3 (c). This is due to the anisotropic feature of single-particle transition around the M point in the reciprocal space, as shown in Fig. 5.5, in which the inter-band transitions with strong oscillator strengths are highly distributed perpendicular to the K-M direction. According to Ref. [30], the optical oscillator strength of exciton states is the summation of the products of the single-particle oscillator strength and electron-hole coupling coefficients $A_{\text{vck}}$ (see Eq. (2.93) and Eq. (2.95) in Chapter 2). Fig. 5.6 schematically shows that this anisotropic feature of single-particle transition leads to a slightly bright \textit{3d} exciton and dark \textit{2p} exciton state. For the \textit{3d} state, the product of the blue-colored $A_{\text{vck}}$ and the single-particle oscillator strength is small but not negligible.
Figure 5.6. (a) Schematic anisotropic feature of single-particle oscillator strength distribution near the M point in the reciprocal space. (b) Schematic electron-hole coupling coefficients distribution of for 2p and 3d exciton states in the reciprocal space, respectively. (c) The overall oscillator strength of 2p and 3d exciton after combination of single-particle oscillator strength and electron-hole coupling coefficients.
Third, the order of these exciton states in Fig. 5.3 (b) is 1s, 2p, 2s, 3d, 3p and 4f, etc., which is qualitatively different from exciton spectra of either 2D or 3D hydrogenic model. Moreover, if we fit the energy dependence of those bright s states according to the main quantum number \( n \), the first-principles result decays much more slowly than the \( \frac{1}{(n-0.5)^2} \) relation derived from the 2D hydrogenic model or the \( \frac{1}{n^2} \) relation derived from the 3D hydrogenic model. These similarities and dissimilarities between \textit{ab initio} results and hydrogenic models encourage us to modify the hydrogenic model by approximating the perpendicular confinement.

### 5.5 Modified Hydrogenic Model for Excitons in 2D Semiconductors

An obvious improvement to the typical 3D hydrogenic model is to confine the Coulomb interaction with a finite width perpendicular to the graphyne layer. In particular, the typical size of excitons shown in Fig. 5.4 is around 10 nm, which is much larger than the thickness of the electron distribution perpendicular to the graphyne plane (\( \sim \) a few Å). This validates the first-order approximation that the thickness of graphyne can be treated as a small number compared to the average distance between electron and hole. As a result, we introduce the following modified Coulomb interaction:

\[
V(r) = -\frac{1}{\varepsilon_0} \frac{1}{\sqrt{r^2 + d_0^2}},
\]

(5.2)
where $r$ is the polar radius of cylindrical coordinates and $d_0$ is the parameter to reflect the effective thickness of 2D excitons. Actually this type of Coulomb interaction had been applied to study many-electron systems before [161, 162].

Then the 2D single-particle Schrödinger equation (in Hartree atomic units) is

$$\left[-\frac{1}{2m^*} \left( \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} \right) - \frac{1}{\sqrt{r^2 + d_0^2}} \right] \cdot \Psi(r, \varphi) = E \cdot \Psi(r, \varphi). \quad (5.3)$$

The effective mass $m^*$ is the reduced mass of electrons and holes ($1/m^* = 1/m_e^* + 1/m_h^*$, averaged by all in-plane directions), which can be obtained by simple DFT calculations because many-electron corrections usually do not change the curvature of electronic bands significantly. The dielectric constant $\varepsilon_0$ is taken as 1 at the first-order approximation for a 2D semiconductor. In a word, only one parameter, the effective thickness $d_0$, is essential in this model.

With the help of the separation of variables $\Psi(r, \varphi) = R(r) \cdot \Phi(\varphi)$, where the angular dependent part is

$$\Phi(\varphi) = \frac{1}{\sqrt{2\pi}} e^{-il\varphi}, l = 0, \pm 1, \pm 2, \ldots \quad (5.4)$$

the radial wavefunction satisfies a 1D Schrödinger equation:

$$\left[\frac{1}{2m^*} \left(-\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{l^2}{r^2} \right) - \frac{1}{\sqrt{r^2 + d_0^2}} \right] \cdot R(r) = E \cdot R(r). \quad (5.5)$$
Eq. (5.5) can be further transformed to a unit-less equation:

\[-\frac{d^2}{dx^2} + \frac{1}{x} \frac{d}{dx} + \frac{l^2 - 1}{x^2} - \frac{2}{\sqrt{x^2 + (m^* \cdot d_0)^2}} \cdot u(x) = \lambda \cdot u(x), \quad (5.6)\]

with boundary condition \(u(x) = 0\) when \(x=0\) and \(u(x) \rightarrow 0\) as \(x \rightarrow \infty\).

All exciton levels can be obtained by solving Eq. (5.6) with a finite-element simulation, and the exciton binding energy levels are

\[E_{nl} = \lambda_{nl} \cdot I \cdot m^*/2, \quad (5.7)\]

where \(I = m_0 e^4/\hbar^2 \approx 27.2 eV\) is the Hartree energy, \(n\) is the principal quantum number and \(l\) is the angular momentum quantum number.

When \(d_0 = 0\), the above model returns to the original 2D hydrogenic model and the analytical solution of exciton binding energy is

\[E_{nl} = \frac{1}{(n - 0.5)^2} \cdot \frac{I \cdot m^*}{2}, \quad n = 1, 2, 3, \ldots \text{ and } l = 0, \pm 1, \pm 2, \ldots, \pm(n - 1). \quad (5.8)\]

In realistic cases, we optimize \(d_0\) according to the energy spacing between the first two bright excitonic (1s and 2s) states, which shall be the easiest data extracted from the optical absorption or luminescence spectrum experiments. In this work, as an example, we fit \(d_0\) according to the energy spacing of 1s and 2s states from the ab initio simulated optical absorption spectrum shown in Fig. 5.3 (a) \((A_1 \text{ and } A_2)\). The results are concluded in Figs. 5.7, in which this modified hydrogenic model provides surpris-
Figure 5.7. Exciton spectra of graphyne from: (a) our model \((m^* = 0.071m_0, d_0 = 2.44\text{nm})\), (b) the ab initio simulation, (c) the original 2D hydrogenic model.

Figure 5.8. Top views of the square of the electron wavefunctions of the lowest-energy excitons from our modified 2D hydrogenic model \((m^* = 0.071m_0, d_0 = 2.44\text{nm})\). The hole is fixed at the center of each plot. The real-space 10-nm scale bars are presented, respectively.
ingly accurate predictions. The deviation of the binding energy between the model and \textit{ab initio} result is less than 40 meV. Considering the extremely light simulation that this requires, this model will be very useful for researchers who are not experts in the field of first-principles MBPT. Besides the binding energies, the eigenstates of the model exhibit exactly the same energy order as those from \textit{ab initio} simulation. More surprisingly, this model even gives the similarly sized wavefunctions for these excitons (see Fig. 5.8) compared to the first-principles results. For comparison purposes, we have also listed the exciton spectrum from the original 2D hydrogenic model in Figs. 5.7, which exhibits substantially larger errors: the binding energy is much smaller than that from \textit{ab initio} simulations; the degeneracy of excitons is not correct and the order of exciton energy levels is qualitatively wrong.

Additionally, we also calculate exciton spectra of another important 2D graphene derivatives, hydrogen-passivated graphene (graphane). The atomistic structure is shown in Fig. 5.9 (a) and the DFT/LDA calculated band structure is plotted in Fig. 5.9 (b). For graphane, two valence bands are degenerate at the $\Gamma$ point, which contribute heavy and light holes simultaneously and thus lead to a mix of low-energy excitons from different holes. Moreover, the lowest-energy exciton of graphane is a relatively dark one because of the charge transfer [134] (shown in Fig. 5.9 (c)), which is qualitatively different from the bright and non-charge-transfer excitons in graphyne. On the other hand, these differences provide us a good opportunity to justify the application range of our model.
Figure 5.9. (a) Top view and side view of the ball-stick model of graphane structure. (b) DFT electronic band structures of graphane. (c) Optical absorption of graphane. The inset is the enlargement of low energy absorption with e-h interaction.
Figure 5.10. Top views of the square of the electron wavefunctions of the characteristic bound excitons in graphane. The hole is fixed at the center of each plot.

Figure 5.11. Exciton spectra of graphane from: (a) our model ($m^* = 0.353m_0, d_0 = 0.54nm$), (b) the \textit{ab initio} simulation, (c) the original 2D hydrogenic model.
Our first-principles GW/BSE calculation identifies a set of low-energy excitons of graphane from the same set of conduction and valence band, meaning we only consider the heavy hole because it contributes to excitons with a larger binding energy. Their real-space wavefunctions are shown in Figs. 5.10. Clearly, these excitons exhibit features of hydrogen-like states. We employ both original 2D hydrogenic model and our modified model, and compare their results with that from \textit{ab initio} simulations, as concluded in Figs. 5.11. We again obtain an excellent agreement between the exciton spectra from our model and \textit{ab initio} simulation. However, the original 2D hydrogenic model still exhibits substantially larger errors and predicts qualitatively wrong results for graphane.

Our model can provide a systematic knowledge of excitons in 2D semiconductors. We have plotted the potential profiles of the bare Coulomb potential and our modified e-h interaction potential in Fig. 5.12 (a). They are significantly different from each other when \( r \) is small, e.g., \( r \) is less than 2 Å. This hints us the most significant corrections from our model is for smaller-sized exciton states. Furthermore, we present how the binding energy of the first three \( s \) exciton states evolves with an effective thickness \( (d_0) \) from the solution of Eq. (5.3), where the effective mass \( m^* = m_0 \) (for other \( m^* \) values, the binding energy and effective thickness are just renormalized by \( m^* \), respectively), in Fig. 5.12 (b). This shows the quantum confinement effects on e-h pairs. For example, we can see the energy spacings between these \( s \) states shrink as we increase the effective thickness \( d_0 \). This explains why graphyne and graphane have a more slowly decaying trend of the exciton binding energies than those of the
original 2D hydrogenic model. Finally, it must be pointed out that our model only works well with 2D semiconductors with a direct band gap, and whose effective masses of electrons and holes are not extremely anisotropic.

![Figure 5.12](image)

Figure 5.12. (a) Bare Coulomb potential and our modified potential for e-h interactions. (b) The evolution of the binding energy of 1s, 2s, and 3s states according to the effective thickness $d_0$ from the solution of Eq. (5.3) when $m^* = m_0$.

5.6 Summary of this Chapter

In summary, we employ the first-principle GW-BSE approach to study optical excitations of graphyne. Our calculation reveals that graphyne is a promising material which may possess the great potential for a wide range of applications, e.g., PV and photo therapy. These quantitative predictions shall be of importance to further the research and application to graphyne. At the same time, we analyze the exciton spec-
tra of graphyne and propose a modified hydrogenic model that explains not only the exciton spectrum of graphyne but also that of graphane, providing a convenient approach for predicting exciton spectra and estimating the binding energies of excitons in 2D semiconductor.
6. ELECTRONIC STRUCTURES AND QUASIPARTICLE ENERGY OF SILICENE STRUCTURES

6.1 Introduction of Silicene

Graphene, a single atomic layer of graphite, is a two-dimensional (2D) allotrope of carbon in hexagonal lattices, as shown in Fig. 6.1 (a). Because graphite is made out of stacks of graphene layers that are weakly coupled by van der Waals (Vd-w) forces, graphene can be produced by the mechanical-exfoliation approach or the micro-mechanical cleavage technique [164]. This was realized experimentally in 2004 [164], and spurred tremendous interest because of the unique linear energy-momentum band dispersion and the subsequent unusual electronic, thermal, and optical properties of graphene. A broad range of exciting applications have been proposed to date [164–168]. For example, the linear energy-momentum dispersion of quasiparticles shown in Fig. 6.1 (b) gives hope to realizing the physics of quantum electrodynamics (QED) of massless Dirac fermions (except for the fact that in graphene the fermions move with a renormalized light speed $v_F \approx 1.1 \times 10^6 m/s$, which is around 300 times smaller than the speed of light $c$ in vacuum) and opening a door for the study of quan-
Figure 6.1. (a) The ball-stick model of graphene structure. The lattice structure is made out of two interpenetrating triangular lattices ($a_1$ and $a_2$ are the lattice unit vectors and $\delta_i \ (i=1,2,3)$ are the nearest-neighbor vectors) (b) Electronic energy dispersion of graphene in reciprocal space. The right panel is an enlargement of the energy bands close to one of the Dirac points.

tum relativistic phenomena in table-top experiments [168]. More generally, graphene represents a conceptually new class of materials that are only one-atom thick, and, on this basis, offers new inroads into low-dimensional physics that never ceases to surprise, and continues to provide a fertile ground for applications [169].

The huge success of graphene has obviously motivated significant efforts to look for the similar 2D honeycomb structures, but made of other group IV elements such as silicon and germanium, which are so called silicene and germanene [170–175, 177–185,
respectively. Primarily because of the current advanced silicon technologies and the strong spin-orbit coupling (SOC) for potential topological insulator candidates [181], silicene and germanene are expected to be promising materials both theoretically and practically.

In realistic experiments, because of the tetrahedral \( sp^3 \) bonding nature of electrons of silicon and germanium atoms, it is extremely challenging to obtain a single layer of silicene or germanene; the intrinsic tetrahedral structure makes silicene/germanene buckled and chemically active. Thus substrates are, therefore, necessary to stabilize the possible layer structures. As a result, unlike the simple planar graphene, depending on the specific substrate, numerous silicene or germanene super structures are proposed by recent experiments with different buckling styles.

The buckling and subsequent variations of atomistic structures are highly like to affect the electronic structures, thus providing a unique degree of freedom to tune the electronic properties of silicene and germanene. In this sense, understanding and further predicting how the change of atomistic structures influences the electronic structures of these novel 2D materials are of central importance for theoretical simulations in this field.

In this chapter, we report first-principles results on the electronic structure of several typical silicene structures. For planar and simply buckled silicenes, we confirm their zero-gap and Dirac-cone nature. Moreover, our simulation reveals a significant renormalization of their Fermi velocities by including self-energy corrections, showing that many-electron effects are crucial for obtaining the electronic and transport
properties of silicene. However, two other recently proposed silicene structures exhibit an unexpected finite band gap, indicating that they are gapped semiconductors instead of previously assumed Dirac-fermion semi-metals. This finite band gap is even preserved after including the Ag substrate. Meanwhile, our GW calculation finds enhanced many-electron effects in these silicene structures because of their reduced dimensionality and, therefore, weaker electronic screening. The GW calculated band gap is in good agreement with experimental measurements. Finally our first-principles simulation proposes that the band gap of the latter two silicene structures can be tuned in a wide range by applying strain, providing a way to engineer electronic and optical properties of these 2D semiconductors.

### 6.2 First-Principles Simulation Setup

Our first-principles calculations employ density functional theory (DFT) and subsequent many-body perturbation theory (MBPT) to reveal the electronic structures and quasiparticle energy of silicene structures. The ground-state calculation is performed by the Quantum Espresso package [79]. For the DFT-level simulations, both local density approximation (LDA) and generalized-gradient approximation (GGA) are applied to check if our conclusions are sensitive to the choice of functionals [4,7]. The Kohn-Sham equation is solved by using a plane-wave basis with a 24 $Ry$ energy cutoff. Norm-conserving pseudopotentials [12] are applied and the k-point sampling grid is $64 \times 64 \times 1$. 

142
Since DFT is known to underestimate band gaps of semiconductors, the single-shot $G_0W_0$ approximation is employed to calculate quasiparticle (QP) band gaps [29,163] with a layered Coulomb truncation of silicene structures. The k-grid sampling of such GW calculations is also conducted on a $64 \times 64 \times 1$ k-grid. The GW calculation is done with the BerkeleyGW package with the general plasmon-pole model [163].

### 6.3 Electronic Structures of Planar and Simply Buckled Silicene Structures

We start from the two most fundamental silicene structures, i.e., the planar and simply buckled ones, as shown in Figs. 6.2 (a) and (b). Unlike graphene, the simply buckled silicene structure is predicted to be more stable than the perfectly planar one [172]. Our first-principles simulation confirms this conclusion; the total energy relation with the buckling distance is presented in Fig. 6.2 (c). When the buckling distance $d$ is zero, the structure converts to the planar one. From Fig. 6.2 (c), we find that the most stable configuration has a buckling distance of 0.4 Å, and its total energy is approximately 20 meV lower than that of planar silicene ($d = 0$). This buckling structure has important impacts on the electronic structure of silicene. For example, recent studies have indicated that the buckling configuration can enlarge the energy band gap by the SOC [181].

Moreover, our DFT calculations reveals that the relaxed planar silicene has a Si-Si bond length of 2.22 Å and that of simply buckled silicene is about 2.20 Å with a
Figure 6.2. The ball-stick model of (a) planar silicene structure and (b) simply buckled silicene structure. The buckling distance is labeled by $d$. For each structure, we present both the top view and side view. The unit cell is illustrated by the shadow region. The different colors of silicon atoms are assigned according to their vertical positions. (c) shows the relation of total energy of one unit cell structure with buckling distance $d$. The energy is shifted so that the one of planar structure is zero.
buckling distance of 0.4 Å, which are in agreement with previous results [172,181,182], confirming the reliability of our DFT simulation.

Here we focus on the electronic structure of silicene around the most interesting region, i.e., the Dirac cone. Figure 6.3 (a) shows the band structures around the Dirac Cone of the simply buckled silicene with and without self-energy correction included. As we expect, the simply buckled silicene exhibits a linear energy-momentum dispersion and a gapless feature. Even after including the many-electron interactions through the GW approximation (the red-colored curve), the gapless feature is preserved while the slope (Fermi velocity) of the band dispersion is substantially increased. For example, it is increased from $v_F = 5.4 \times 10^5$ m/s at the DFT level to $7.4 \times 10^5$ m/s by the GW calculation. This is evidenced by Fig. 6.3 (b), in which the quasiparticle energy and LDA band energy fall into a single linear line. Thus

Figure 6.3. (Color online) (a) DFT and GW calculated band structure around the Dirac cone of simply buckled silicene. (b) The quasiparticle energy vs. LDA band energy of simply buckled silicene, which shows a linear correction with a slope of 1.37.
the many-electron correction is also linear and continuous, maintaining the zero-gap band dispersion. Interestingly, we find the planar silicene has a nearly identical band structure as that of the simply buckled one. The Fermi velocity of the planar silicene is \( v_F = 5.6 \times 10^5 \text{ m/s} \) at the DFT-level simulation and is increased to \( 7.7 \times 10^5 \text{ m/s} \) by the GW calculation. These Fermi velocities of both simply buckled and planar silicene structures are substantially smaller than that of graphene (\( \sim 1.1 \times 10^6 \text{ m/s} \)) but the renormalization of the Fermi velocity by self-energy corrections is similar to that of graphene, which is around a 37% enhancement from their DFT result because of the depressed screening effect and enhanced electron-electron interactions [189,190].

In summary, both planar and simply buckled silicene structures always possess a linear band dispersion and a gapless character, which are independent on the choice of calculation methods. The fundamental reason is that the AB sublattice symmetry of these structures, which protects the linear band crossing.

### 6.4 A Brief Review of Recently Proposed Silicene Structures

As discussed in the introduction, substrates are crucial for stabilizing the silicene structure. Because of different choices of substrates, a number of new silicene superstructures have been proposed according to the layer-substrate interaction. Meanwhile, because of the complicated substrate effects, many experimental results and interpretations are highly controversial.
For instance, in a recent publication by Lalmi et al., the synthesis 2D silicene on the Ag(111) surface was claimed based on scanning tunneling microscopy (STM) observations [175]. However, the silicene sheet determined by Lalmi et al. is strongly compressed with a Si-Si distance of only 0.19 nm (± 0.01 nm), a value far smaller compared to the value of 0.23 nm expected from the DFT calculations by Cheng et al. [176] and our own study. Because no other complementary experimental results have been presented, the conclusion of this silicene synthesis is still regarded to be rather speculative. Moreover, the STM results of Lalmi et al. for the Si/Ag(111) structure have not been reproduceable by other experimental groups [177, 178, 182]. Thus it is finally believed that the measurements presented by Lalmi et al. refer to the clean Ag(111) surface itself because the pure Ag(111)-(1×1) surface can mimic a honeycomb-like appearance in STM caused by a tip-induced contrast reversal [177], which looks very similar to the STM images reported by Lalmi et al.

Recently, Vogt et al. reported the synthesis of epitaxial silicene sheets on the Ag(111) substrate through the combination of STM and angular-resolved photoemission spectroscopy (ARPES) [182]. They proposed a (4×4) silicene structure, as shown in Figure 6.4 (a). In this structure, a unit cell of the (3×3) silicene layer is adjusted to match the (4×4) Ag(111) surface lattice. Thus the (4×4) notation is labeled with respect to the Ag lattice for this type of structure. The DFT simulated image of this proposed structure exhibits the same structural features as those observed in the experimental STM image, shown in Figure 6.4 (b).
Figure 6.4. (a) Top view and side view of the fully relaxed atomic geometries of the model for silicene on the Ag(111) surface; right side is the enlarged image of the hexagonal silicene ring indicated by the white circle in left. (b) Simulated STM image (left) for the structure shown in (a) and the observed STM image (right) in the experiment. (c) ARPES intensity map for the clean Ag surface (left) and after formation of the 2D Si adlayer (right), taken along the Ag $\bar{\Gamma}$-$\bar{K}$ direction through the silicene $\bar{K}$ ($h\nu = 126$ eV). Right side is the Brillouin-zone (BZ) scheme of the 2D Si layer with respect to the Ag(111)-(1 × 1) surface. The red arrow indicates the ARPES measurement direction. [182]
According to the ARPES measurement of the silicene layer on the Ag (111) substrate, a linear dispersion of energy-momentum relation is observed in Figure 6.4 (c), with a Fermi velocity of $v_F = 1.3 \times 10^6$ m/s, which is higher than the one found for graphene [188]. Although there is a finite band gap shown in their measurement (Figure 6.4 (c)), they attributed it to the substrate-layer interaction and thus claimed to find a novel $(4 \times 4)$ silicene structure with a massless Dirac Fermion dispersion and zero-gap nature.

Shortly, Chen et al. proposed another silicene structure [184]. To explain their observation of silicene on the Ag(111) substrate by STM and spectroscopy studies [184], they propose a $(\sqrt{3} \times \sqrt{3})$ structure that has a double-side buckling. The details of the structure is presented in Figure 6.5 (a) and (b), in which the red Si atoms are buckled upward and the green ones are buckled downward. It has to be pointed out that this name of $\sqrt{3} \times \sqrt{3}$ is different from the definition of the $(4 \times 4)$ structure because here this structure is labeled with respect to the intrinsic silicene $(1 \times 1)$ unit cell instead of the Ag substrate. We keep using these names to follow the mainstream of experimental literatures in this field.

Moreover, for the $(\sqrt{3} \times \sqrt{3})$ structure, Chen et al. claim that its quasiparticles behave as massless Dirac fermions because a linear energy-momentum dispersion is derived from quasiparticle interferences (QPI) patterns, as shown in Figure 6.5 (c).

More results on silicene super structures have been published recently [183, 185–187, 200, 201]. These works reported the $(\sqrt{3} \times \sqrt{3})$ phase [185], the $(4 \times 4)$ phase, [200] and the rhombic $(\sqrt{3} \times \sqrt{3}R30^\circ)$ phase [201] on the Ag(111) surface. A $(\sqrt{3} \times$
Figure 6.5. (a) The top view and side view of the lattice geometry of the double-buckled $(\sqrt{3} \times \sqrt{3})$ silicene superstructure. (b) A larger schematic model illuminating the honeycomb structure of $(\sqrt{3} \times \sqrt{3})$ reconstructed silicene. (c) Energy dispersion as a function of $\kappa$ for silicene determined from the wavelength of QPI patterns. The inset shows a schematic drawing of the overall band structure, with the relative location of $E_F$, and the data points [red (thick gray) line]. [184]
phase on the ZrB2 substrate [183] was discovered as well. However, there is a fundamental question that has not been solved yet, i.e., whether these silicene structures still preserve the original Dirac-fermion characteristics, provided that their structures are so different to each other.

Unfortunately, very limited attempts have been focused on the electronic structures of these newly proposed silicene structures, which is necessary for guiding experimental measurements and understanding the electronic and optical properties of these increasingly popular materials. In particular, the complicated buckling styles may break the symmetry of the honeycomb lattices. According to the history of graphene, the variation of structures and symmetries will not only change the Fermi velocity but also possibly generate a finite band gap. This may prove to be a crucial advantage over graphene for broader applications, such as bipolar devices and high-performance field-effect transistors (FETs).

6.5 Electronic Structures of Experimentally Proposed Silicene Structures

Here we choose two typical phases, the $(\sqrt{3} \times \sqrt{3})$ and $(4 \times 4)$ structures, proposed by Vogt et al. [182] and Chen et al. [184] respectively. Our motivation is to reveal the how these different buckling styles modify the electronic structure of these novel materials.

The atomistic configurations of these two structures are shown in Figure 6.6 (a) and (b). The main differences between them are the size of the unit cell and the way to buckle the lattices. For example, the silicon atoms in the $(\sqrt{3} \times \sqrt{3})$ structure
shown in Fig. 6.6 (a) are located in three horizontal planes while those in the (4 × 4) structure shown in Fig. 6.6 (b) are located in two horizontal planes.

![Figure 6.6](image_url)

Figure 6.6. The ball-stick model of experimentally proposed silicene structures. (a) buckled silicene with a \((\sqrt{3} \times \sqrt{3})\) unit cell; (b) buckled silicene with a \((4 \times 4)\) unit cell. The buckling distance is labeled by \(d\). For each structure, we present both the top view and side view. The unit cell is illustrated by the shadow region. The different colors of silicon atoms are assigned according to their vertical positions.

It has to be pointed out that those suspended structures shown in Figs. 6.6 (a) and (b) do not exactly mimic the experimental cases, in which the substrate may be an essential factor [182, 184]. This is indicated by our DFT calculations and other published works showing that these isolated structures are only metastable [184]. Therefore, as what have been done before [194], our solution is to fix the buckling distance \(d\) and relax all other degrees of freedom, including force and stress.
Our following calculations will show that the essential physical picture and main conclusions shall not be affected by this approximation.

![Figure 6.7](image)

Figure 6.7. (a) Band structures of the $\sqrt{3} \times \sqrt{3}$ silicene structures with buckling distance (i) $d = 0$ Å, (ii) $d = 0.48$ Å, (iii) $d = 0.71$ Å, (iv) $d = 1.17$ Å. (b) Band structures of the $4 \times 4$ silicene structures with buckling distance (i) $d = 0$ Å, (ii) $d = 0.46$ Å, (iii) $d = 0.75$ Å, (iv) $d = 1.09$ Å. The top of valence band is always set to be zero.
The band structures and its evolution to the buckling distance $d$ of the \((\sqrt{3} \times \sqrt{3})\) silicene structure are shown in Fig. 6.7 (a). Initially, for the buckling distance $d = 0$ Å, the silicene is planar and the Dirac cone is at the $\Gamma$ point due to the band folding. However, when the buckling distance is increased, there appears a finite band gap at the $\Gamma$ point. As the buckling distance $d$ further increases, the gap increases and reaches about 0.21 eV when the buckling distance $d = 0.71$ Å, which is the typical buckling value measured from experiments [172,182,183,185]. Therefore, we surprisingly conclude that the Dirac-fermion character is no longer preserved in this \((\sqrt{3} \times \sqrt{3})\) structure. This is qualitatively different from the previously claimed massless Dirac-fermion dispersion by Ref. [184]. Actually, we can obtain a well-defined effective mass of the free carriers from the curvature of the band dispersion in Fig. 6.7 (a) (iii). For example, the nearly isotropic effective mass of the light electron is 0.083 $m_0$ and that of the hole is 0.079 $m_0$, which are typical values of semiconductors.

We further tune the buckling distance $d$ while maintaining the supercell geometry to obtain the evolution of the band gap, which is presented in Fig. 6.8 (a). From another point of view, this is equivalent to tuning the strain condition of the layer structure. For both LDA and GGA results, the band gap of such a \((\sqrt{3} \times \sqrt{3})\) structure increases similarly.

When the buckling distance is more than 0.8 Å, we see a transition of the band gap from the direct one to the indirect one, which is reasonable because the band structure shall approach that of tetrahedral silicon due to the stronger $sp^3$ hybridization. The
corresponding band structure is shown in Fig. 6.7 (a) (IV), in which the conduction band minimum is shifted along the $\Gamma - M$ direction.

To further confirm this finite band-gap nature, we have performed the GW calculation on the ($\sqrt{3} \times \sqrt{3}$) structure. For the purpose of justification, one GW calculation of a typical buckling distance shall be enough; as shown in Fig. 6.8 (a), the QP band gap is around 0.5 eV for the $d = 0.71$ Å buckled case, which is almost a 140% enhancement from the DFT result because of the substantially depressed screening [124,149]. It must be addressed again that our referenced experiments use a metallic substrate, which may reduce the self-energy correction due to the metallic screening and possible charge transfer. However, all of these factors will not close the finite band gap and, usually, the realistic QP band gap shall still be slightly larger than the DFT result [191,192].

We now turn to the band structures of the other important ($4 \times 4$) silicene phase, as shown in Fig. 6.7 (b). This time we observe a finite band gap again for the nonzero buckling distance. When the buckling distance is $d = 0.75$ Å that is chosen from an experimental study [182], the ($4 \times 4$) structure is a typical direct-band-gap semiconductor as shown in Fig. 6.7 (b)(iii). The evolution of the band gap with the buckling distance $d$ is presented in Fig. 6.8 (b) as well. Similar to the ($\sqrt{3} \times \sqrt{3}$) case, the band gap increases as we enlarge the buckling distance. When the buckling distance is more than $d = 0.8$ Å, the similar direct-to-indirect band gap transition is observed in Fig. 6.8 (b). The corresponding GW results are marked here, predicting a larger QP band gap than the DFT results.
Interestingly, when we consult the relevant experimental results, a finite band gap was actually observed there, but it had been attributed to extrinsic factors, such as the layer-substrate interaction [182] or substrate-induced symmetry breaking [200]. However, our simulation shows that this finite band gap may be intrinsic if the silicene sample possesses their claimed $(4 \times 4)$ structure. Moreover, if we compare the experimentally observed band gap value by the ARPES to our first-principles GW result shown in Fig. 6.8 (b), they are very close to each other, around 0.6 eV [182]. This consistency does not necessarily mean that our simplified model exactly mimics experimental conditions because many extrinsic factors have not been included in our simulation. More work is needed to verify the quantitative band gap in silicene.

![Figure 6.8. The band gap evolution as the variation of buckling distance $d$ for (a) the $(\sqrt{3} \times \sqrt{3})$ silicene structure and (b) the $(4 \times 4)$ silicene structure. The direct band gap values are connected by dash lines while the indirect band gap values are connected by solid lines. The GW results are marked by cross signs.](image)
After presenting the above first-principles results, it is necessary to figure out the physical reason for such a finite band-gap opening in these promising silicene structures. According to past intensive studies on graphene, a necessary condition for the massless Dirac cone with a zero-gap character is the inversion symmetry of AB sublattices. However, for the structures shown in Figs. 6.6 (a) and (b), A and B sublattices are no longer equivalent to each other, resulting in the broken inversion symmetry and, thereafter, generating a finite band gap. A similar idea was noticed before, e.g., the applied gating electric field can break this symmetry and introduce a finite band gap in simply buckled silicene and germanene. [193–195].

![Figure 6.9. Top view of color contour plots of the charge distribution of typical conduction and valence states at energy extrema around the band gap for (a) the planar silicene when $d = 0$ and (b) the $(4 \times 4)$ silicene structure when $d = 0.75 \text{ Å}$.

157
In order to better see this broken AB sublattice symmetry, we present the charge distributions of valence and conduction states at energy extrema around the band gap. As shown in Fig. 6.9 (a), for planar silicene, we observe the charge distributions of both valence and conduction states are nearly identical for A and B sublattices. However, for the (4 × 4) silicene structure, we can clearly see the inequivalent charge distributions of A and B sublattices as shown in Fig. 6.9 (b) due to the broken symmetry.

All the above calculations consider only the free-standing silicene structures. However, most experimentally achieved silicenes are on metallic substrates. To include substrate effects, we have additionally calculated the (√3 × √3) and (4 × 4) silicene structures on top of three Ag (111) substrate layers. Because of the Vdw-interaction nature between silicene and the substrate, we employ the Vdw functional implemented in the Vienna \textit{ab initio} simulation package (VASP) [197–199] to relax atomistic structures and obtain the electronic structure. The silicene-substrate distance is chosen by minimizing the total energy of the system. For example, in Fig. 6.10 (b), the optimized distance is found to be about 6.0 Å for the (4 × 4) silicene structure on the Ag (111) surface.

Because the large number of atoms in such a big supercell results in an extremely messy band structure, it is impossible to directly tell which bands are from silicene or silver. Therefore, we present the projected density of states (PDOS) of silicon and silver, respectively, to identify the band gap of silicene. In Fig. 6.11 (a1), for the isolated (√3 × √3) silicene, we see the band gap is around 0.20 eV. In Fig. 6.11 (a2),
regarding this ($\sqrt{3} \times \sqrt{3}$) silicene on the Ag substrate, its PDOS of silicon is nearly the same as the isolated case and the energy gap is around 0.15 eV, being slightly smaller due to the layer-substrate interaction and corresponding minor structure changes. Similarly, for the $(4 \times 4)$ silicene, the Ag substrate reduces the band gap slightly and the PDOS illustrates that a gap of about 0.20 eV still exists, as shown in Figs. 6.11 (b1) and (b2).

![Figure 6.10](image)

Figure 6.10. The total energy relation with the distance between the $(4 \times 4)$ silicene and Ag (111) substrate (a) without Vdw interaction, and (b) with Vdw interaction.

This Vdw calculation further confirms that the recently proposed silicene structures have qualitatively different electronic structures and band gaps from the widely assumed massless Dirac-fermion dispersion by previous studies [182, 184]. The most likely reason for this conflict between our simulation results and previous experimental conclusions may be due to the interpretations of atomistic structures from their
Figure 6.11. (a1) DOS of the isolated \((\sqrt{3} \times \sqrt{3})\) silicene \((d = 0.71 \text{ Å})\). (a2) PDOS of the \((\sqrt{3} \times \sqrt{3})\) silicene on the Ag(111) substrate. (b1) DOS of the isolated \((4 \times 4)\) silicene \((d = 0.75 \text{ Å})\). (b2) PDOS of the \((4 \times 4)\) silicene on the Ag(111) substrate. The Fermi level is marked by the dashed line, respectively.
experimental data because many extrinsic factors can impact the STM image, misleading their explanations. On the other hand, our calculation does not include some factors that may give rise to such unusual band structures, such as the SOC. Therefore, we expect further experimental and theoretical work to reveal new physics in these novel silicene structures.

Finally we have to point out that the finite-gap nature of these recently proposed silicene structures shall be of practical interest. As shown in Figs. 6.7 and 6.8, the band gap and even the band topology can be tuned in a wide range by the buckling distance $d$ and the associated strain; this provides a precious degree of freedom to control the electronic and optical properties of such materials, making them superior to gapless graphene. According to previous studies [172,175,182,183,185], the lattice structure and buckling distance are able to be varied by layer-substrate interactions. Therefore, if future experiments can fabricate and confirm more similar structures, they will be of broad interest for the development of devices that employ a tunable band gap.

During the reviewing process of our work by Applied Physics Letters, results for a finite band gap of silicene sheets induced by broken symmetry were published [200,201]. Lin’s work [200] reported the $(4 \times 4)$ silicene on the Ag (111) surface loses its Dirac fermion characteristics, and Chen’s work [201] reported a gap opening in their rhombic $(\sqrt{3} \times \sqrt{3}R30^\circ)$ silicene on the Ag(111) surface. Their work confirm our prediction that the $(\sqrt{3} \times \sqrt{3})$ and $(4 \times 4)$ silicene structures don’t preserve Dirac fermions and have a gap opening. However, while Lin and Chen attribute them to the
broken symmetry from the substrate, we conclude that they result from the broken symmetry of the buckled silicene structures themselves.

6.6 Summary of this Chapter

We employ first-principles simulations to study the electronic structure of various silicenes. Our results show that the band structure of silicene is sensitive to the buckled atomistic structure. Unlike the planar and simply buckled ones, those silicene structures proposed by recent experiments exhibit a finite band gap, making them promising 2D semiconductors instead of zero-gap semimetals, as were previously assumed. This conclusion is further confirmed by including the Vdw interaction with the Ag substrate. On the other hand, this finite band gap can be tuned in a wide range by strains, making these silicene structures superior to graphene for many important semiconductor applications. Our predictions reinterpret key experimental measurements and will motivate more reliable justifications.
7. CONCLUSIONS

7.1 Conclusions of Projects

We have employed the first-principles simulation to study the electronic structures, lattice vibrational modes, and optical excitations of important reduced-dimensional semiconductors, including the one-dimensional (1D) Si/Ge core-shell nanowires (NWs), the two-dimensional (2D) graphene derivatives and silicene.

First, we present how to engineer the band alignment of Si/Ge core-shell NWs. Our calculations reveal that band alignments in Si/Ge core-shell NWs are significantly modified by the strain condition and quantum confinement. By applying axial strain, the valence band offset can be efficiently tuned across a wide range and even be diminished. We find two mechanisms that contribute to the physical origins of the strain effects of band offsets. Furthermore, varying the size of Si/Ge core-shell NWs and corresponding quantum confinement is shown to be efficient for tuning both valence and conduction band offsets simultaneously. Our proposed approaches to control band offsets in nano-sized heterojunctions may be of practical interest for nanoelectronic and photovoltaic applications.
Second, we study the lattice vibrational modes of Si/Ge core-shell NWs. Our calculations show that the internal strain induced by the lattice mismatch between core and shell plays an important role in significantly shifting the frequency of characteristic optical modes of core-shell NWs. In particular, our simulation demonstrates that these frequency shifts can be detected by Raman-scattering experiments, giving rise to a convenient and nondestructive way to obtain structural information of core-shell materials. Meanwhile, another type of collective modes, the radial breathing modes (RBM), is identified in Si-core/Ge-shell NWs and their frequency dependence is explained by an elastic media model. Our studied vibrational modes and their frequency evolution are useful for thermoelectric applications based on core-shell nanostructures.

Third, we studied optical properties and exciton spectra of 2D semiconducting carbon structures. The energy spectra and wavefunctions of excitons in the 2D graphene derivatives, i.e., graphyne and graphane, are found to be strongly modified by anisotropic quantum confinement, making them qualitatively different from the usual Rydberg series. We propose a one-parameter hydrogenic model to explain the ab initio exciton spectra, and it yields a quantitatively good agreement with the results from first-principles calculations. Therefore, our model provides a convenient way to extrapolate the electron-hole binding energy from optical spectroscopies of 2D semiconductors without needing costly simulations. Meanwhile, the high optical absorbance and large spin singlet-triplet splitting make graphyne a promising candidate for energy and biomedical applications.
Lastly, we reported electronic structure calculations of 2D graphene-like systems, i.e., silicene. For planar and simply buckled silicene structures, we confirm their zero-gap nature and show a significant renormalization of their Fermi velocity by including many-electron effects. However, the other two more complicated buckled silicene structures proposed by recent experiments exhibit a finite band gap, indicating that they are gapped semiconductors instead of expected Dirac-fermion semimetals. The finite band gap nature of the latter two structures is preserved even with the Ag substrate included. The gap opening is explained by the inversion symmetry breaking of the buckled structures. Moreover, our GW calculation reveals enhanced many-electron effects in these 2D structures. Finally the band gap of the latter two structures can be tuned in a wide range by applying strain.

7.2 Future Work

Based on the above studies, there are several new projects that can be done as future works.

For example, the investigation of excitons in bilayer 2D semiconductors could be very interesting since the electron-hole pairs may be coupled not only within a single layer but also between the layers. In bilayer hexagonal boron nitride (BN) or bilayer graphene, we expect that the features of exciton spectra and selection rules of optical transitions may be different from what we observe in single-layer systems. Moreover, different ways of stacking bilayer structures can provide different inter-layer coupling and interactions, as people have observed in bilayer graphene.
In particular, another unique degree of freedom is the twisting of bilayer structures. From our preliminary calculations of optical absorptions of bilayer BN, twisted-structures yield significant different optical spectra from those of the AA and AB stacking cases. For instance, a new peak appears below the dominant peak. How to understand these new properties induced by the twisting and the underlying physics shall be a very interesting project.
REFERENCES


VITA

Shouting Huang was born in a small town of southern China. He obtained a B.S. degree from University of Science and Technology of China in 2007. Since August 2007, he has been a teaching/research assistant in Department of Physics, Washington University in St. Louis. In April 2010, he joined Dr. Li Yang’s group. His research focuses on first-principles calculations of electronic, optical, and thermal properties of reduced-dimensional materials. His post-graduation plan is to work as a seismic imaging engineer/researcher in Houston, Texas.

Publications:


