博士論文

First-Principles Study on Photoenergy Conversion in Semiconductors using Density-functional and Many-body Perturbation Theories

（密度汎関数理論及び多体摂動理論を用いた半導体における光エネルギー変換に関する第一原理的研究）

河合　宏樹

First-Principles Study on Photoenergy Conversion in Semiconductors using Density-functional and Many-body Perturbation Theories

Hiroki Kawai

Department of Chemical System Engineering

Graduate School of Engineering

The University of Tokyo

February 2015

Preface

Theoretical study is a fundamental and necessary step towards the full understanding of the physical properties of materials. In particular, analyses based on quantum mechanics are required to fully disclose electronic properties of materials. Nowadays, along with the great advances of quantum chemistry, solid state physics, and computer science, numerical calculation of electronic properties is able to provide accurate results without any empirical parameter (in *first-principles* manner). Accordingly, such computational approach, integrating experimental studies, is becoming a powerful tool to reveal the microscopic nature of materials.

This thesis summarizes computational studies on semiconductors with enhanced photoenergy conversion capability: gallium zinc oxynitride (GaN:ZnO), zinc-blende gallium nitride (*zb*-GaN), and organic-inorganic/inorganic lead halide perovskite ($APbI_3$). Such materials are receiving deep attention due to their extremely appealing features in photoenergy conversion. Their properties related to photoabsorption and carrier diffusion have been investigated by means of state-of-the-art first-principles approaches: density functional theory (DFT) and many-body perturbation theory (MBPT).

The outline of this thesis is as follows: Chapter 1 is the general introduction for this thesis. Description of photoenergy conversion systems, i.e., photovoltaics and photocatalysis, and the

important physical properties for such photoenergy conversion systems, i.e., bandgap, band bending, and carrier lifetime, are presented. Motivations for the adopted computational approach are also provided.

In Chapter 2, theoretical backgrounds of DFT and MBPT are described. DFT is well known and widespread in the community of quantum chemistry and solid state physics (Walter Kohn and John Pople, Nobel Prize in Chemistry 1998). However, this method is applicable only for ground state. Although MBPT is one of the most successful approaches to correct DFT shortcomings for calculating excited-state properties, it is still not widely spread in the community of theoretical chemistry. In this chapter, I tried to give straightforward descriptions of MBPT for readers unfamiliar with this theory.

In Chapter 3, studies of bandgap and band bending are presented focusing on GaN:ZnO, which is an overall water-splitting photocatalytic material with visible-light response. The mechanism of bandgap narrowing by alloying GaN and ZnO is studied by calculating bandgaps and band edge positions for many alloy structures. Furthermore, band bending on the surface of GaN:ZnO due to an electric double layer stemming from donor-rich and acceptor-rich layers is predicted.

In Chapter 4, studies on the effects of electron-phonon interaction on optical absorption and carrier lifetime are presented focusing on two semiconductors: *zb*-GaN and $APbI_3$. For the former system, electron-phonon interaction renormalizes bandgap and introduces band

broadening, which results in bandgap and optical absorption spectrum perfectly reproducing the experimental result. For the latter, the mechanism of relaxation of hot carriers by electron-phonon interaction is discussed. My results clearly explain the mechanism of the experimentally reported slow hot hole cooling in $CH_3NH_3PbI_3$, a material with a key role in perovskite solar cells.

In Chapter 5, concluding remarks and future outlooks are presented.

Here, I would like to show acknowledgements for people supporting my work. At first, I thank Professor Koichi Yamashita, my supervisor. Since I joined to his group in 2009, he has provided me many opportunities to get in touch with various kinds of scientific communities ranging from domestic experimentalists to foreign theoreticians. Such experiences largely opened my mind. I also thank Associate Professor Hiroshi Ushiyama for his encouragement and cares in my daily life.

Most of works in this thesis are joined with Dr. Giacomo Giorgi. Since I was an undergraduate student, he always helped me with comprehensive mind. He introduced me to international scientific activity. With his great cares, my two visits to Roma in 2011 and 2012/2013 were really fruitful. It is definitely one of the happiest in my life to have such a kind Italian friend sharing the scientific interests with me. I express my immeasurable gratitude here.

I have also involved several Italian researchers. In my first visit to Roma in 2011, Dr. Maurizia Palummo in University of Rome “Tor Vergata” kindly received me and taught me MBPT. Even though my stay was only one month, I still vividly remember the days in Tor Vergata. I would like to thank her. In my second visit in 2012/2013, Dr. Andrea Marini in Consiglio Nazionale delle Ricerche (CNR), who is the father of YAMBO code, kindly accepted my request to learn electron-phonon interactions in MBPT scheme with him. Discussions with him were really stimulating for me. I always move to his pure personality and faithful attitude to physics. I would like to acknowledge his great help to my work.

I thank my colleagues in Yamashita-Ushiyama group. Dr. Mikiya Fujii gave me many opportunities to consider physics and a grounding to be a member of society. His great effort allowed me to concentrate on scientific activity. Dr. Ryota Jono helped me a lot since I was an undergraduate student. His sincerity and philosophy to the science positively affected my mind. Discussion with Ms. Eriko Watanabe about photocatalysis and electrochemistry largely improved my activities in my PhD days. Mr. Tomoyuki Hata and Mr. Katsuhiko Nishimura also helped me to improve my knowledge, especially about quantum mechanics and photovoltaics, in daily discussions. Mr. Tomotaka Kunisada and Ms. Ayaka Kuroki stimulated me as PhD students in the same grade. Two secretaries, Ms. Mayumi Iyama and Ms. Noriko Sanpei, kindly supported my activities. I thank Dr. Giorgi, Ms. Watanabe, Mr. Hata, and Dr. Jono again for their careful checks of my thesis.

I’m also grateful to the former members of Yamashita-Ushiyama group. Dr. Hideyuki Kamisaka, who is a project assistant professor in the University of Tokyo, and Dr. Hisao Nakamura, who is a senior researcher in National Institute of Advanced Industrial Science and Technology, provided me a lot of suggestions from technical and scientific point of view. Dr. Tatsuhiko Ohto, who is a research associate in Osaka University, initiated me to study hard especially about solid state physics, Green’s function technique, and electron-phonon interactions. His helps definitely widened my scope. I also thank Dr. Yu Xuefang, Dr. Jong-Won Song, Dr. Yasumitsu Suzuki, Mr. Shotarou Sudo, Dr. Takaya Ogawa, Mr. Kosuke Nakai, and Mr. Yusuke Kudo for sharing an enjoyable time with me.

Finally, I want to thank my family for the fundamental supports.

Tokyo　February 2015　　*Hiroki Kawai*

Contents

Chapter 1.
General Introduction

1-1. Photovoltaics and photocatalysis

In this 21st century, seven billion of people populating the earth necessitate an ever increasing amount of energy. Needless to say, further developments of human society are extremely conditioned by the availability of electricity and chemical fuel from heat, light, wind, fossil, and nuclear, via safe and clean manufacturing processes. Among the various forms of energy, sunlight is most abundant and ubiquitous. Solar energy shining surface of the earth is estimated ~25 TJ/km^{2}day (in the assumption that photon flux is 1 kW/m^2 and 7 hours sunshine duration in a day). In principle, such huge amount of energy is sufficient to fully cover human needs.

Photovoltaics (hereafter, PVs) and photosynthesis are photoenergy conversion systems into electricity and chemical fuels, respectively. Direct photoenergy conversion into electricity using PVs is efficient and hence solar cells are widely employed in the world. As well as PVs, water-splitting photocatalysis has also received considerable attention as artificial photosynthesis of hydrogen. Hydrogen represents a promising clean energy carrier because it

does not produce any greenhouse gases or polluting materials when transformed into heat via combustion and electricity via fuel cell. One of the main methods for hydrogen production is steam reforming of natural gas, which emits carbon dioxides. Therefore photocatalytic water splitting is expected to be an alternative clean, viable, and environmentally friendly manufacturing process of hydrogen production. In both PVs and photocatalysis, the basic processes of photoenergy conversion have been and still are subject of deep analysis in the scientific community to achieve higher energy conversion efficiencies.

Development of PV device and photocatalyst stems from the discovery of the photovoltaic effect, which was described by Becquerel for the first time in 1839.[1] His photoelectrochemical experiment demonstrated that a voltage and an electric current are produced when a silver chloride electrode is dipped into an electrolyte and illuminated. The mechanism of this phenomenon, after called “Becquerel effect”, was not well understood at that time. In the first half of twentieth century, quantum mechanics had been developed starting from light quantum hypothesis proposed by Einstein and the concept of photon was accepted. In the latter half of twentieth century, it was revealed that the mechanism of the Becquerel effect is generation of electron-hole pairs in a semiconductor by photon absorption and their effective separation due to the electric field formed at the solid-electrolyte interface (space charge layer).

The first photovoltaic cell was realized in Bell laboratory by Chapin, Fuller, and Pearson in 1954,[2] more than one century after the initial studies of Becquerel. Assembling silicon *p-n* junction cell they observed that electron-hole pairs generated by irradiation are separated and make to do work in an external circuit. Nowadays, silicon solar cells using monocrystalline, polycrystalline, and amorphous silicon are prevalent.[3]

For further dissemination of solar cells, breakthroughs to bring reduction of the fabrication cost and improvement of the power conversion efficiencies (PCEs) are strongly desired. Accordingly, new classes of PVs, e.g., thin-film copper indium gallium selenide (CIGS) solar cell, dye-synthesized solar cell (DSSC)[4], organic PV (OPV)[5] devices, quantum-dot sensitized solar cell[6], and perovskite solar cell[7], are emerging and attracting attention. Compared with standard silicon solar cells, such new PVs have a number of merits from the viewpoint of cost, flexibility, and property tunability. According to diversification of the PVs, many kinds of semiconductor materials are potential candidates as photoabsorbers. To improve the PCEs and the durability of these new classes of PVs, physical properties of a large number of organic molecules, inorganic chemical compounds, inorganic nanoparticles, and organic-inorganic hybrid materials have been extensively investigated.

The first photocatalytic water-splitting reaction was observed in 1972 on photoelectrochemical cell using a photoanode of TiO_2 and a photocathode of Pt. In this cell, TiO_2 absorbs photons and then photogenerated holes cause oxygen evolution reaction (OER)

on the surface of TiO_2. At the same time, photogenerated electrons flow to an external circuit and then hydrogen evolution reaction (HER) occurs on the Pt electrode. This is well known as Honda-Fujishima effect.[8] In 1977, Nozik proposed to directly couple the photoanode of semiconductor and the photocathode of metal without an electric circuit (Schottky-type photochemical diode).[9] As one form of such photochemical diodes, the composite of powders of a semiconductor and nanoparticles of a metal is preferable with respect to the scalability and hence widely investigated. In this system, the semiconductor is called the "photocatalyst" and the metal is called the "cocatalyst".[10]

In the past three decades, many semiconductors have been explored for searching more and more efficient water-splitting photocatalysts: oxides of transition metals with empty d-orbitals (d^0-type) such as $NaTaO_3$ and typical metals with filled d-orbitals (d^{10}-type) such as β-Ga_2O_3.[10] These metal oxides are, however, not suitable for visible-light absorption due to the too wide bandgaps, thus oxide doping and/or their sensitization have been investigated. A two-step water-splitting reaction, i.e., Z-scheme, is also attracting much interest. Non-oxide such as (oxy)nitride and oxysulfide have also been extensively studied.

As described above, many semiconducting materials are being explored both for PV device and photocatalyst.

1-2. Basic processes of photoenergy conversion

For improving the efficiency of photoenergy conversion, deep understanding of the basic mechanism that rules the conversion process becomes mandatory. The initial processes are common in both PVs and photocatalysis: the first process is the photoabsorption that generates electron-hole pairs in the absorber, while the second is the diffusion of carriers towards the interface/surface of the absorber. After the diffusion, the carriers are transferred out of the absorber, specifically toward carrier transport layers in the case of PV device, toward a cocatalyst metal or liquid in the case of photocatalytic process.

In this subsection I will describe the details of the initial two processes and introduce the important physical properties that rule these processes.

1) Photoabsorption

Photoabsorption results in excitation of the absorber. In molecular orbital theory and band theory, an excitation is regarded as a transition of electrons lying in the occupied orbitals, i.e., valence band, to the unoccupied orbitals, i.e., conduction band. Here, the bandgap (E_g) is defined as

$$E_g = E_{\mathrm{CBM}} - E_{\mathrm{VBM}}, \qquad (1.1)$$

where E_{CBM} and E_{VBM} are energies of the conduction band minimum (CBM) and of the valence band maximum (VBM). E_g is the “electronic gap” (or “quasiparticle gap”) neglecting electron-hole interaction in excited states. When E_g is narrowed by the exciton

binding energy it becomes the “optical gap”, which corresponds to the experimental optical absorption edge. However, in the case of inorganic compounds, the electronic gap is still a good measurement of photoabsorption since the exciton binding energies are usually small due to the screening effect, e.g., Wannier-Mott excitons.

E_g governs the efficiency of photoenergy conversion from following two main key points. The first point is the number of photogenerated carriers. Photogenerated carriers are decreased as E_g becomes larger since no photon is absorbed when the energy of photon $h\nu$ is lower than E_g (h and ν are Planck’ constant and frequency of photon, respectively). Indeed, the spectrum of the solar radiation has strong intensity in the visible-light range from 380 nm (~3.3 eV) to 780 nm (~1.6 eV), thus gaps smaller than 3 eV are preferable to gain a large number of photogenerated carriers.

The second is the potential of photogenerated carriers. In the radiation with energies larger than the bandgap ($h\nu > E_g$), “hot” carriers, i.e., electrons with higher energy than E_{CBM} (hot electron) and holes with lower energy than E_{VBM} (hot hole), are generated. However, they rapidly decay towards the band edges via scattering with other carriers, impurities, and phonons (thermalization). Thus, in general, the potentials of excited electrons and holes in photon irradiation are determined by E_{CBM} and E_{VBM}. It means that the excess energy $\Delta E = h\nu - E_g$ is usually lost in the process of photoenergy conversion.

Accordingly, E_g must be optimized to maximize the number of carriers and minimize energy loss ΔE corresponding to the solar spectrum. In PVs, the electric power is a product of an electric current and a voltage, which are determined by the number of photogenerated carriers and the potentials of excited electrons and holes in the absorber, respectively. Shockley and Queisser suggested optimal E_g is 1.34 eV in single *p-n* junction cell to maximize the PCE.[11]

In the case of photocatalyis, the carrier's potential determines the ability of redox reactions: electrons must have more negative potential than the redox potential of the desired reduction reaction, and similarly, potentials of holes must be more positive than the redox potential of the oxidation reaction. Otherwise, even if a number of carriers are generated, they do not result in redox reactions at all. Concerning the water-splitting photocatalyst, the minimal requirement is $E_g = 1.23$ eV since the redox potential of HER

$$2\mathrm{H}^+ + 2\mathrm{e}^- \rightarrow \mathrm{H}_2, \qquad (1.2)$$

is 0 V (vs. NHE) and that of OER

$$2\mathrm{H}_2\mathrm{O} \rightarrow \mathrm{O}_2 + 4\mathrm{H}^+ + 4\mathrm{e}^-. \qquad (1.3)$$

is +1.23 V (vs. NHE). In practice, however, gaps wider than 1.23 eV are required in order to overcome the activation barrier (overpotential) for charge transfer between the photocatalyst and the water molecules.[10] Hence most of the photocatalysts achieving overall water splitting have bandgaps in ultraviolet wavelengths.

2) Carrier diffusion

The carrier diffusion is the process that takes place after the photoabsorption and described by the carrier diffusion length, L defined as[12]

$$L = \sqrt{D\tau^{rec}}, \tag{1.4}$$

where D is the diffusion constant and τ^{rec} is the carrier lifetime of electron-hole recombination. Diffusion lengths of electron and hole are required to be sufficiently long to reach the outside of the absorber. Diffusion constant is related with the mobility, μ in the Einstein relation as

$$D = \frac{\mu k_B T}{q}, \tag{1.5}$$

where k_B is the Boltzmann constant, T is the temperature, and q is the charge of the carrier. Furthermore, μ is defined as

$$\mu = \frac{e\tau^{coll}}{m^*}, \tag{1.6}$$

where e, τ^{coll}, and m^* are the elementary charge, the carrier collision time, and the carrier effective masses. From Eq (1.4), (1.5), and (1.6), L will clearly be prolonged by long τ^{rec}, long τ^{coll}, and small m^*.

Band bending, as space charge layer at *p*-*n* junction, promotes charge separation and hence effectively extends recombination lifetime τ^{rec}. Reducing trap density in the bulk and surface of the absorber also results in long τ^{rec}. Collision lifetime τ^{coll} is determined by how often carriers are scattered with electrons, phonons, and impurities. Effective mass m^*

is related to the wavefunction of valence and conduction bands: m^* of hole (electron) is smaller as the valence (conduction) band is delocalized.

From the above discussions, I can conclude that it is mandatory to optimize bandgap, band bending, carrier lifetime, localization of band edges, and trap density in an absorber material for achieving effective photoenergy conversion. Works in this thesis focus on and discuss these properties, especially bandgap, band bending, and carrier lifetime.

1-3. Theoretical study on photoenergy conversion process

To “talor” materials with desired properties of photoabsorption and carrier diffusion, an improved knowledge of the relationship between the chemical structures of materials and the physical properties is required. The properties mentioned in the previous subsection all depend on the electronic structure of absorber materials. Thus, once Schrödinger equation is solved on the electronic Hamiltonian constituted by the atomic structures of chemical compounds, we can predict the physical properties. However, deriving the exact solution of such *first-principles* Schrödinger equation is impossible for many-body systems, thus some approximations must be introduced.

One of the most widely accepted approximation is Kohn-Sham approach based on density functional theory (DFT). The complicated many-body problem turns into an accessible self-consistent solution of a set of single-particle equations (Kohn-Sham equations). DFT calculations employing local density approximation (LDA) and their derivatives on the exchange correlation functional are surprisingly successful, and thus popular, in order to describe the electronic properties of atoms, molecules, and solids.

Despite an undeniable success in the prediction of structural and thermodynamic properties of materials, DFT remains a theory for the description of ground-state properties, thus not applicable to excited states. In other words, it does not provide correct band energies for empty states leading to errors in bandgap predictions. To correct the error,

electron-electron interaction must be described in a more accurate formalism. Moreover, as DFT is the theory for electronic Hamiltonian, it does not tell us anything about electron-hole and electron-phonon interactions, which are important to understand optical bandgap and carrier decay process.

Many-body perturbation theory (MBPT) is one of state-of-the-art theories to overcome the DFT drawbacks. In particular, MBPT introduces the interactions not described at DFT level such as correct electron-electron and electron-hole interactions, which are necessary to describe excited-state properties. Methods to treat electron-electron and electro-hole interactions are called GW approximation and Bethe-Salpeter equations, respectively. MBPT also provides theoretical treatment of electron-phonon interaction, which is usually neglected in the Born-Oppenheimer approximation which separates electronic motions from nuclear motions. Considering electron-phonon interaction, it is possible to compute carrier lifetimes and discuss the relaxation process. Due to the development of computational resources and simulation packages, DFT combined with MBPT is becoming a powerful tool in material science.

Virtue of theoretical study is in the modeling. In computational studies, we can start from the simplest ideal model, e.g. pure chemical compounds, single molecules, clean surfaces, and then these models can be extend to more complicated ones, e.g. defective crystals, supermolecules, solid solutions, amorphous, interfaces, etc. Therefore, we can address the

effects of the complicated factors individually, which is opposite to experimental studies that directly analyze complicated systems in nature. Combining these opposite approaches, i.e., *bottom-up* computational approach and *top-down* experimental approach, understanding of the electronic processes is deepened from theoretical point of view even in complicated systems, e.g. metal oxynitride, organic-inorganic hybrid material.

Bibliography

(1) Williams, R. *J. Chem. Phys.* **1960**, *32*, 1505.

(2) Chapin, D. M.; Fuller, C. S.; Pearson, G. L. *J. Appl. Phys.* **1954**, *25*, 676.

(3) Goetzberger, A.; Hebling, C.; Schock, H.-W. *Mater. Sci. Eng. R Reports* **2003**, *40*, 1.

(4) Grätzel, M. *Acc. Chem. Res.* **2009**, *42*, 1788.

(5) Hoppe, H.; Sariciftci, N. S. *J. Mater. Res.* **2011**, *19*, 1924.

(6) Rühle, S.; Shalom, M.; Zaban, A. *Chemphyschem* **2010**, *11*, 2290.

(7) Green, M. a.; Ho-Baillie, A.; Snaith, H. J. *Nat. Photonics* **2014**, *8*, 506.

(8) Fujishima, A.; Honda, K. *Nature* **1972**, *238*, 37.

(9) Nozik, A. J. *Appl. Phys. Lett.* **1977**, *30*, 567.

(10) Maeda, K. *J. Photochem. Photobiol. C Photochem. Rev.* **2011**, *12*, 237.

(11) Shockley, W.; Queisser, H. J. *J. Appl. Phys.* **1961**, *32*, 510.

(12) Ashcroft, N. W.; Mermin, D. N. *Solid state physics*; Brooks/Cole Cengage Learning: Australia, 1976.

Chapter 2.
Density-Functional Theory and Many-Body Perturbation Theory

2-1. Density Functional Theory

If we assume the Born-Oppenheimer approximation, i.e. nuclear motions negligible compared to electronic ones, the electronic Hamiltonian, $\hat{H}$ is defined by the number of electrons, N, the atomic species, Z_I, the number of ions, N_I, and the atomic positions, $\mathbf{R}_I$, as

$$\begin{aligned}\hat{H} &= \hat{T} + \hat{U} + \hat{V}_{\text{ext}} \\ &= -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N}\sum_{I=1}^{N_I}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}\end{aligned} \tag{2.1}$$

where $\hat{T}$ is the kinetic energy of electrons, $\hat{U}$ is the Coulomb interaction between electrons, and $\hat{V}_{\text{ext}}$ is the interaction between electrons and ions. Here atomic unit ($m_e = \hbar = e = 1$) is used. Solving the Schrödinger equation $\hat{H}|\Psi\rangle = E|\Psi\rangle$ is impossible for many-body systems.

Density functional theory (DFT) is a theory that enables to solve quantum many-body problems for the ground state. Nowadays, DFT represents, *de facto*, the standard approach to calculate electronic structure of atoms, molecules, and solids. In this section, I will briefly introduce the basic concept of DFT.[1]

2-1-1. Hohenberg-Kohn theorem

In 1964, Hohenberg and Kohn formulated the theorems[2] that define the ground-state energy *E* as a universal functional in terms of ground-state electronic density $n_0(\mathbf{r})$ and the correct ground-state energy as a result of $E[n_0(\mathbf{r})]$ minimization. In details,

Theorem 1

$V_{\text{ext}}(\mathbf{r})$ is uniquely determined by the density of the ground state $n_0(\mathbf{r})$ apart from a trivial additive constant.

To demonstrate this theorem, we assume there are two external potentials $V_{ext}^1(\mathbf{r})$ and $V_{ext}^2(\mathbf{r})$ that provide the ground-state electronic density $n_0(\mathbf{r})$. Accordingly, two Hamiltonians $\hat{H}^1$ and $\hat{H}^2$ are determined resulting in two ground states $|\Psi_0^1\rangle$ and $|\Psi_0^2\rangle$.

$|\Psi_0^2\rangle$ is not the ground state for $\hat{H}^1$, thus we can see

$$\begin{aligned} E_1 = \langle\Psi_0^1|\hat{H}^1|\Psi_0^1\rangle &< \langle\Psi_0^2|\hat{H}^1|\Psi_0^2\rangle \\ &= \langle\Psi_0^2|\hat{H}^2|\Psi_0^2\rangle + \langle\Psi_0^2|\hat{V}_{ext}^1 - \hat{V}_{ext}^2|\Psi_0^2\rangle \qquad (2.2) \\ &= E_2 + \int d\mathbf{r}[V_{ext}^1(\mathbf{r}) - V_{ext}^2(\mathbf{r})]n_0(\mathbf{r}). \end{aligned}$$

Similarly, the following relation is also fulfilled

$$E_2 < E_1 + \int d\mathbf{r}[V_{ext}^2(\mathbf{r}) - V_{ext}^1(\mathbf{r})]n_0(\mathbf{r}). \qquad (2.3)$$

Adding Eq (2.2) to (2.3) leads to the inconsistency,

$$E_1 + E_2 < E_2 + E_1. \tag{2.4}$$

Thus, $V_{\text{ext}}(\mathbf{r})$ is uniquely determined by the density of the ground state $n_0(\mathbf{r})$ apart from a trivial additive constant.

Solving the Schrödinger equation with the Hamiltonian including $\hat{V}_{ext}$, the wavefunction in the ground state $|\Psi_0\rangle$ is uniquely determined in the case of non-degenerate ground state. Thus, once $\hat{V}_{ext}$ is given, the electronic density $n_0 = \langle\Psi_0|\Psi_0\rangle$ is determined. This theorem shows that the ground-state energy *E* is a unique functional of the electronic density $n_0(\mathbf{r})$, $E[n_0(\mathbf{r})]$.

Since $|\Psi_0\rangle$ is functional of the density in the ground-state $n_0(\mathbf{r})$, the ground-state total energy is defined as

$$E[n_0(\mathbf{r})] = F[n_0(\mathbf{r})] + \int d\mathbf{r} V_{ext}(\mathbf{r}) n_0(\mathbf{r}), \tag{2.5}$$

where $F[n_0(\mathbf{r})]$ is an *universal* functional of the kinetic energy and the Coulomb interaction

$$F[n_0(\mathbf{r})] = \langle\Psi_0|\hat{T} + \hat{U}|\Psi_0\rangle. \tag{2.6}$$

Theorem 2

The exact ground-state energy is determined as the global minimum value of $E[n_0(\mathbf{r})]$ for any particular $\hat{V}_{ext}$ and the density minimizing $E[n_0(\mathbf{r})]$ is the exact ground-state density $n_0^{\text{exact}}(\mathbf{r})$.

To prove this second theorem, we recall that since $E[n_0(\mathbf{r})]$ is uniquely determined by the ground-state wavefunction $|\Psi_0\rangle$ the variational principle leads to the following relation,

$$E[n_0^{exact}(\boldsymbol{r})] = \langle \Psi_0^{exact} | \hat{H} | \Psi_0^{exact} \rangle < \langle \Psi_0' | \hat{H} | \Psi_0' \rangle = E[n_0'(\boldsymbol{r})], \qquad (2.7)$$

thus $E[n_0(\mathbf{r})]$ is minimized when $n_0(\mathbf{r}) = n_0^{\text{exact}}(\mathbf{r})$.

The initial proof of Hohenberg and Kohn is restricted to densities that are in the ground state under some potential $\hat{V}_{ext}$. Such densities are called "*V*-representable". However, it is not clarified which conditions are required for densities to satisfy the *V*-representability. Levy and Lieb have extended the Hohenberg-Kohn theorem to "*N*-representable" densities[3,4], i.e. densities derived from any wavefunctions for *N* electrons. The conditions that the "*N*-representable" densities should satisfy are: 1) $n_0(\mathbf{r}) > 0$ and 2) $\int \left|\nabla {n_0}^{1/2}\right|^2 (\mathbf{r})$ is finite.

2-1-2. Kohn-Sham equation

Although Hohenberg-Kohn theorems offer the proof that the ground-state energy is obtained by minimizing $E[n(\mathbf{r})]$, the exact form of the universal functional $F[n(\mathbf{r})]$ is not provided. The *ansatz* of Kohn and Sham[5] is that the density of original interacting particles is equal to that of the chosen virtual independent particles;

$$n(\mathbf{r}) = \sum_{\sigma} \sum_{i=1}^{N^{\sigma}} |\psi_i^{\sigma}(\mathbf{r})|^2, \tag{2.8}$$

where $\{\psi_i^{\sigma}(\mathbf{r})\}$ are wavefunctions of independent particles called Kohn-Sham (KS) orbitals and N^{σ} is the number of electrons with spin σ.

By introducing the kinetic energy of KS orbitals, T_{s}, given by

$$T_{\mathrm{s}}[n] = -\frac{1}{2} \sum_{\sigma} \sum_{i=1}^{N^{\sigma}} \langle \psi_i^{\sigma} | \nabla^2 | \psi_i^{\sigma} \rangle, \tag{2.9}$$

and the Hartree term of classical Coulomb interaction E_{Hartree} defined as

$$E_{\mathrm{Hartree}}[n] = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{2.10}$$

the energy functional in Eq (2.5) is reformulated as

$$E_{\mathrm{KS}}[n] = T_{\mathrm{s}}[n] + E_{\mathrm{Hartree}}[n] + E_{\mathrm{XC}}[n] + \int d\mathbf{r} V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}), \tag{2.11}$$

where $E_{\mathrm{XC}}[n]$ is known as the exchange-correlational function which includes all many-body effects,

$$E_{\mathrm{XC}}[n] = F[n] - T_{\mathrm{s}}[n] - E_{\mathrm{Hartree}}[n]. \tag{2.12}$$

Eq (2.11) indicates that ground-state energies of many-body systems can be obtained once the

exact form of $E_{\mathrm{XC}}[n]$ is known. Unfortunately, the exact form of $E_{\mathrm{XC}}[n]$ is unknown, but some practical approximate functionals provide successful results. One of such approximations is the LDA that will be presented in the following section.

Solution of Eq (2.11) is given by the Lagrange multiplier method with the orthonormalization constraints on KS orbitals $\langle \psi_i^{\sigma} | \psi_j^{\sigma'} \rangle = \delta_{i,j}\delta_{\sigma,\sigma'}$;

$$\frac{\delta}{\delta \psi_i^{\sigma *}(\mathbf{r})}\left[E_{\mathrm{KS}}[n] - \sum_{j,\sigma'} \varepsilon_j^{\sigma'}\left(\langle \psi_j^{\sigma'} | \psi_j^{\sigma'} \rangle - 1\right)\right] = 0, \tag{2.13}$$

where $\varepsilon_j^{\sigma'}$ is a multiplier. In the four terms in the right-hand sides of Eq (2.11), only T_{s} is explicitly expressed as a functional of KS orbitals. Thus, using the chain rule, Eq (2.13) can be written as

$$\begin{aligned} &-\frac{1}{2}\nabla^2 \psi_i^{\sigma}(\mathbf{r}) + \left[\frac{\delta E_{\mathrm{Hartree}}[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{XC}}[n]}{\delta n(\mathbf{r})} + V_{\mathrm{ext}}(\mathbf{r})\right]\frac{\delta n(\mathbf{r})}{\delta \psi_i^{\sigma *}} \\ &= \left[-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r})\right]\psi_i^{\sigma}(\mathbf{r}) = \varepsilon_i^{\sigma}\psi_i^{\sigma}(\mathbf{r}). \end{aligned} \tag{2.14}$$

This Schrödinger-like set of equations is known as the Kohn-Sham (KS) equations. The effective potential $V_{\mathrm{eff}}(\mathbf{r})$ depends on the density $n(\mathbf{r}')$,

$$V_{\mathrm{eff}}(\mathbf{r}) = \frac{1}{2}\int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{\mathrm{XC}}[n]}{\delta n(\mathbf{r})} + V_{\mathrm{ext}}(\mathbf{r}), \tag{2.15}$$

thus, Eq (2.8), (2.14) and (2.15) must be solved self-consistently.

2-1-3. Local Density Approximation

The relevance of KS approach is that the kinetic energies of independent particles and the long-range Hartree terms are clearly separated from the universal functional $F[n]$. We can assume that the remaining exchange-correlational functional $E_{XC}[n]$ can be reasonably approximated as a local or nearly local functional of the density as

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})\varepsilon_{\mathrm{XC}}([n], \mathbf{r}), \tag{2.16}$$

where $\epsilon_{\mathrm{XC}}([n], \mathbf{r})$ is the exchange-correlation energy per electron at **r**, which is determined by the density $n(\mathbf{r})$ close to **r**. This is the local density approximation (LDA). LDA is exact in a homogeneous electron gas since the exchange and correlation energies are characterized by the constant density. Since $E_{\mathrm{XC}}[n]$ is universal, the exact form of $\varepsilon_{\mathrm{XC}}([n], \mathbf{r})$ can be obtained by calculating a homogeneous electron gas.

The exact exchange energy of a homogeneous electron gas provided by the Hartree-Fock approximation is known as

$$\varepsilon_{\mathrm{X}}[n] = -\frac{3}{4}\left(\frac{3n}{\pi}\right)^{1/3}, \tag{2.17}$$

in the unpolarized system. It is straightforward to expand Eq (2.17) to the polarized case. By the total density $n = n^{\uparrow} + n^{\downarrow}$ and the fractional polarization,

$$\xi = \frac{n^{\uparrow} - n^{\downarrow}}{n}, \tag{2.18}$$

the exchange part in a polarized system $\varepsilon_{\mathrm{X}}[n, \xi]$ is written as

$$\varepsilon_{\mathrm{X}}[n, \xi] = \varepsilon_{\mathrm{X}}[n, 0] + \left[\varepsilon_{\mathrm{X}}[n, 1] - \varepsilon_{\mathrm{X}}[n, 0]\right] f_{\mathrm{x}}(\xi), \tag{2.19}$$

where $f_{\mathrm{x}}(\xi)$ is defined as follows,

$$f_{\mathrm{x}}(\xi) = \frac{1}{2}\frac{(1+\xi)^{4/3} + (1-\xi)^{4/3} - 2}{2^{1/3} - 1}. \tag{2.20}$$

In contrast to the case of exchange, the analytic form is not obtained for the correlation functional $\varepsilon_{\mathrm{C}}[n]$. Thus the functional is constructed by fitting the numerically calculated correlation energies by quantum Monte Carlo methods over a wide range of densities. The forms of the functional and its parameterization are shown in, for example, Ref [6,7].

Starting from LDA, generalized gradient approximation (GGA) has been developed to include inhomogeneous density.[8,9] This functional also depends on the derivatives of density as

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})\varepsilon_{\mathrm{XC}}(n(\mathbf{r}), |\nabla n(\mathbf{r})|). \tag{2.21}$$

The advent of GGA has boosted the use of DFT in quantum chemistry. In particular, GGA functional parameterized by Perdew-Burke-Ernzerhof (PBE)[10] is one of the most widely used.

As further improvement, hybrid functional combining Hartree-Fock (HF) exchange with explicit density functional have been shown to better reproduce many structural/electronic properties of several chemical systems. PBE0 functional is constituted by the correlation energy of PBE and the exchange of HF with 25% and that of PBE with 75%.[11] Heyd, Scuseria, and Ernzerhof (HSE) also have introduced functional similar to PBE0, which has exchange of HF with 25% only in the short range (screened exchange).[12,13]

2-2. Many-Body Perturbation Theory

KS-DFT along with LDA/GGA functional provides successful results for the ground-state property description for atoms, molecules, and solids in terms of their electronic structures and geometry by the Hellman-Feynman forces. However, it might fail to describe bandgaps because the effective potentials of LDA/GGA do not satisfactorily reproduce the response to change of electronic distribution under excitation processes.

Many-body perturbation theory (MBPT) takes advantage of the usage of technique based on the Green's function in order to consider many-body interaction as a perturbation to non-interacting systems. Once we consider the electron-electron interaction as a perturbation, dynamical and non-local effective potential, i.e., the self-energy, is obtained by GW approximation (Sec 2-3). As a result, a *quasiparticles* band structure comparable with that resulting from photoemission spectroscopy is obtained.

Furthermore, a perturbation scheme of polarizability is also formulated. Bethe-Salpeter equation (BSE) is one of the approximations for polarizability, which provides excitonic states that includes electron-hole (e-h) interactions (Sec 2-4). It enables us to consider excited states so that computational results can be compared with photoabsorption experiments.

In this section, theoretical background of MBPT, which is fundamental in order to understand GW approximation and BSE, is summarized based on the descriptions in Ref [14–17].

2-2-1. Green's function

To treat interactions as a perturbation on non-interacting system, Hamiltonian $\hat{H}$ is separated into non-interacting Hamiltonian $\hat{H}_0$ and an interaction $\hat{V}$,

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{2.22}$$

In the case of electronic total Hamiltonian shown in Eq (2.1), electron-electron interaction is regarded as a perturbation $\hat{V} = \hat{U}$ into non-interacting system $\hat{H}_0 = \hat{T} + \hat{V}_{\text{ext}}$.

In second quantization the field operators $\hat{\psi}(x)$ and $\hat{\psi}^\dagger(x)$ are defined as

$$\hat{\psi}(x) = \sum_\lambda \varphi_\lambda(x) c_\lambda, \quad \hat{\psi}^\dagger(x) = \sum_\lambda \varphi_\lambda(x) c_\lambda^\dagger \tag{2.23}$$

where c_λ and $c_\lambda^\dagger$ are Fermion's creation and annihilation operators, respectively, on state λ with wavefunction $\varphi_\lambda(x)$. These operators satisfy anticommutator rule as

$$\{c_\lambda, c_{\lambda'}^\dagger\} = c_\lambda c_{\lambda'}^\dagger + c_{\lambda'}^\dagger c_\lambda = \delta_{\lambda,\lambda'}, \tag{2.24}$$

$$\{c_\lambda^\dagger, c_{\lambda'}^\dagger\} = \{c_\lambda, c_{\lambda'}\} = 0. \tag{2.25}$$

and similarly, field operators satisfy the rule

$$\{\hat{\psi}(x), \hat{\psi}^\dagger(x')\} = \delta(x - x'), \tag{2.26}$$

$$\{\hat{\psi}^\dagger(x), \hat{\psi}^\dagger(x')\} = \{\hat{\psi}(x), \hat{\psi}(x')\} = 0. \tag{2.27}$$

A time-ordered Green's function at zero temperature is defined as

$$G(x,t;x',t') = -i\frac{\langle \Psi_0 | T\{\hat{\psi}_{\text{H}}(x,t)\hat{\psi}_{\text{H}}^\dagger(x',t')\} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}, \tag{2.28}$$

where $\hat{\psi}_{\text{H}}(x,t) = e^{iHt}\hat{\psi}(x,0)e^{-iHt}$ and $\hat{\psi}_{\text{H}}^\dagger(x',t') = e^{iHt'}\,\hat{\psi}^\dagger(x',0)e^{-iHt'}$. Subscript H indicates that the operator is represented in the Heisenberg picture. $|\Psi_0\rangle$ is the ground-state

of the interacting system $\hat{H}|\Psi_0\rangle = E|\Psi_0\rangle$. $T\{\cdots\}$ is the time-ordering operator (T-product) which moves the operator in earlier time to the right;

$$T\{A(t)B(t')\} = \theta(t-t')A(t)B(t') \mp \theta(t'-t)B(t')A(t). \quad (2.29)$$

The sign of the second term on the right-hand side is determined by the number of Fermion operators in $A(t)$ and $B(t')$. Due to the anticommutation rule shown in Eq (2.26), the sign is negative when $A(t) = \hat{\psi}_{\mathrm{H}}(x,t)$ and $B(t') = \hat{\psi}_{\mathrm{H}}^{\dagger}(x',t')$.

If one differentiates $G(x,t;x',t')$ with respect to the time *t*, the equation of motion

$$i\frac{\partial}{\partial t}G(x,t;x',t') = \delta(t-t')\delta(x-x') - i\frac{\left\langle \Psi_0 \middle| T\left\{ i\frac{\partial}{\partial t}\hat{\psi}_{\mathrm{H}}(x,t)\hat{\psi}_{\mathrm{H}}^{\dagger}(x',t') \right\} \middle| \Psi_0 \right\rangle}{\langle \Psi_0|\Psi_0\rangle}, \quad (2.30)$$

is obtained. The first term on the right-hand side is obtained by the derivative of the Heaviside functions in Eq (2.29) and the relation of Eq (2.26). For example, single-particle system with the Hamiltonian $\hat{H} = \hat{H}_0$ provides $i\frac{\partial}{\partial t}\hat{\psi}_{\mathrm{H}}(x,t) = \left\{\hat{\psi}_{\mathrm{H}}(x,t), \hat{H}_0\right\} = H_0(x)\hat{\psi}_{\mathrm{H}}(x,t)$, thus Eq (2.30) can be written as

$$\left[i\frac{\partial}{\partial t} - H_0(x) \right] G_0(x,t;x',t') = \delta(t-t')\delta(x-x'), \quad (2.31)$$

where $G_0(x,t;x',t')$ is the non-interacting Green's function. Eq (2.31) clearly shows that $G_0(x,t;x',t')$ is Green's function for time-dependent Schrödinger equation.

To solve the many-body problem, the Green's function for fully interacting Hamiltonian is required. Such fully interacting Green's function is obtained as perturbations to the non-interacting Green's function G_0.

Non-interacting Green's function can be calculated as follows.

$$\begin{aligned} & G_0^{\mathbf{p}\sigma}(t,t') \\ &= -i\sum_{\mathbf{p},\sigma}\langle\Phi_0|T\{c_{\mathbf{p}\sigma}(t)c_{\mathbf{p}\sigma}^{\dagger}(t')\}|\Phi_0\rangle \\ &= -i\sum_{\mathbf{p},\sigma}\{\theta(t-t')\langle\Phi_0|c_{\mathbf{p}\sigma}(t)c_{\mathbf{p}\sigma}^{\dagger}(t')|\Phi_0\rangle - \theta(t'-t)\langle\Phi_0|c_{\mathbf{p}\sigma}^{\dagger}(t')c_{\mathbf{p}\sigma}(t)|\Phi_0\rangle\}, \end{aligned} \tag{2.32}$$

where $|\Phi_0\rangle$ is the ground state of the non-interacting system, i.e., $\hat{H}_0|\Phi_0\rangle = E|\Phi_0\rangle$. Here state λ is defined by momentum and spin index $\lambda = (\mathbf{p},\sigma)$.

$c_{\mathbf{p}\sigma}(t) = e^{iH_0t}c_{\mathbf{p}\sigma}e^{-iH_0t}$ is transformed by Taylor expansion as

$$c_{\mathbf{p}\sigma}(t) = c_{\mathbf{p}\sigma} + [iH_0t, c_{\mathbf{p}\sigma}] + \frac{1}{2!}[iH_0, [iH_0t, c_{\mathbf{p}\sigma}]] + \cdots. \tag{2.33}$$

Since $[iH_0, c_{\mathbf{p}\sigma}] = -iH_0c_{\mathbf{p}\sigma}$ for non-interacting Hamiltonian, the previous equation is transformed into

$$\begin{aligned} c_{\mathbf{p}\sigma}(t) &= c_{\mathbf{p}\sigma} - iH_0tc_{\mathbf{p}\sigma} + \frac{1}{2!}(-iH_0t)^2c_{\mathbf{p}\sigma} + \cdots \\ &= \mathrm{e}^{-i\varepsilon_{\mathbf{p}\sigma}t}c_{\mathbf{p}\sigma}, \end{aligned} \tag{2.34}$$

where $H_0 = \sum_{\mathbf{p}\sigma}\varepsilon_{\mathbf{p}\sigma}c_{\mathbf{p}\sigma}^{\dagger}c_{\mathbf{p}\sigma}$. Similarly, we can also derive

$$c_{\mathbf{p}\sigma}^{\dagger}(t) = \mathrm{e}^{i\varepsilon_{\mathbf{p}\sigma}t}c_{\mathbf{p}\sigma}^{\dagger}. \tag{2.35}$$

In the ground state, the occupancy of state $\lambda = (\mathbf{p},\sigma)$ is ruled by the Fermi-Dirac distribution function $f_{\mathbf{p}\sigma}$,

$$\begin{aligned} \langle\Phi_0|c_{\mathbf{p}\sigma}^{\dagger}c_{\mathbf{p}\sigma}|\Phi_0\rangle &= f_{\mathbf{p}\sigma} \\ \langle\Phi_0|c_{\mathbf{p}\sigma}c_{\mathbf{p}\sigma}^{\dagger}|\Phi_0\rangle &= 1 - f_{\mathbf{p}\sigma}. \end{aligned} \tag{2.36}$$

Substituting Eq (2.34), (2.35), and (2.36) into Eq (2.32),

$$G_0^{\mathbf{p},\sigma}(t-t') = -i\mathrm{e}^{-i\varepsilon_{\mathbf{p}\sigma}(t-t')}\{\theta(t-t')(1-f_{\mathbf{p}\sigma}) - \theta(t'-t)f_{\mathbf{p}\sigma}\}. \tag{2.37}$$

The Fourier transform of $G_0^{\mathbf{p},\sigma}(t)$ provides

$$\begin{aligned} G_0^{\mathbf{p},\sigma}(\omega) &= -i\left[\int_0^{\infty} dt\mathrm{e}^{i(\omega-\varepsilon_{\mathbf{p}\sigma}+i\delta)t}(1-f_{\mathbf{p}\sigma}) - \int_{-\infty}^{0} dt\mathrm{e}^{i(\omega-\varepsilon_{\mathbf{p}\sigma}-i\delta)t}f_{\mathbf{p}\sigma}\right] \\ &= \frac{1-f_{\mathbf{p}\sigma}}{\omega-\varepsilon_{\mathbf{p}\sigma}+i\delta} + \frac{f_{\mathbf{p}\sigma}}{\omega-\varepsilon_{\mathbf{p}\sigma}-i\delta}, \end{aligned} \tag{2.38}$$

where the infinitesimal quantity δ is introduced to get the integral convergence.

2-2-2. Perturbative expansion of the Green's function

In the definition of *G* in Eq (2.28), the exact ground state of interacting system $|\Psi_0\rangle$ is included, but yet unknown. To treat *G* in a perturbative way, we have to link $|\Psi_0\rangle$ with the ground state of non-interacting system $|\Phi_0\rangle$. This is obtained rewriting Eq (2.28) in the interaction picture.

Three representation pictures of quantum mechanics are summarized below.

1) Schrödinger picture

   The wavefunctions are time-dependent: $\psi(t) = e^{-iHt}\psi(0)$;the operators are time-independent.

2) Heisenberg picture

   The wavefunctions are time-independent; the operators are time-dependent: $\hat{A}(t) = e^{iHt}\hat{A}(0)e^{-iHt}$.

3) Interaction pictures

   The wavefunctions are developed under perturbative interaction in the Hamiltonian $\hat{H}(t) = \hat{H}_0(t) + \hat{V}$: $\psi(t) = e^{iH_0t}e^{-iHt}\psi(0)$; the operators developed under non-interacting Hamiltonian: $\hat{A}(t) = e^{iH_0t}\hat{A}(0)e^{-iH_0t}$.

At first, we define *U*-matrix as below;

$$U(t) = e^{iH_0 t} e^{-iHt}, \quad \hat{\psi}(t) = U(t)\hat{\psi}(0). \tag{2.39}$$

Note that we assume the case of $[\hat{H}_0, \hat{V}] \neq 0$, thus $U(t) \neq e^{-iVt}$. Equation of motion of $U(t)$ is written as

$$\begin{aligned} \frac{\partial}{\partial t} U(t) &= i\left[e^{i\hat{H}_0 t}\hat{H}_0 e^{-i\hat{H}t} - e^{i\hat{H}_0 t}\hat{H}e^{-i\hat{H}t}\right] \\ &= ie^{i\hat{H}_0 t}\left(\hat{H}_0 - \hat{H}\right)e^{-i\hat{H}t} \\ &= -ie^{i\hat{H}_0 t}\hat{V}e^{-i\hat{H}t} \\ &= -ie^{i\hat{H}_0 t}\hat{V}e^{-i\hat{H}_0 t}U(t) \\ &= -i\hat{V}(t)U(t), \end{aligned} \tag{2.40}$$

then we get

$$U(t) - U(0) = -i\int_0^t \mathrm{d}t_1 \hat{V}(t_1)U(t_1). \tag{2.41}$$

Rearranging it with $U(0) = 1$ gives

$$U(t) = 1 - i\int_0^t \mathrm{d}t_1 \hat{V}(t_1)U(t_1), \tag{2.42}$$

and by iteratively substituting this equation into *U* on the right-hand side, we get

$$U(t) = 1 + \sum_{n=1}^{\infty}(-i)^n \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \cdots \int_0^{t_{n-1}} \mathrm{d}t_n \hat{V}(t_1)\hat{V}(t_2)\cdots\hat{V}(t_n). \tag{2.43}$$

Furthermore, this equation can be rewritten using *T*-product introduced by Eq (2.29) as

$$\begin{aligned} U(t) &= 1 + \sum_{n=1}^{\infty}\frac{(-i)^n}{n!}\int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \cdots \int_0^t \mathrm{d}t_n T\left\{\hat{V}(t_1)\hat{V}(t_2)\cdots\hat{V}(t_n)\right\} \\ &= T\exp\left(-i\int_0^t \mathrm{d}t_1 \hat{V}(t_1)\right). \end{aligned} \tag{2.44}$$

This relation is easily confirmed by checking the case of n = 2.

From U-matrix, S-matrix can be defined as

$$S(t,t') = U(t)U^{\dagger}(t') \quad \hat{\psi}(t) = S(t,t')\hat{\psi}(t'), \tag{2.45}$$

and the solution is also written in the same manner as Eq (2.44);

$$S(t,t') = T\exp\left(-i\int_{t'}^{t} \mathrm{d}t_1 \hat{V}(t_1)\right) \tag{2.46}$$

The development of wavefunction in interaction picture is written using S-matrix as $\psi(t) = S(t,0)\psi(0)$. Due to the relation of $S(0,t)S(t,0) = 1$, it is evident that

$$\psi(0) = S(0,t)\psi(t). \tag{2.47}$$

If we assume that the system feels no interaction at $t \to -\infty$ and the interaction is introduced moving toward $t = 0$ adiabatically, the exact ground states $\psi(0)$ and those in non-interacting system Φ_0 are related as

$$\psi(0) = S(0,-\infty)\Phi_0. \tag{2.48}$$

This relationship is known as Gell-Mann and Low theorem. The wavefunctions in Heisenberg picture $|\Psi_0\rangle$ is identical with one in the interaction picture at $t = 0$, thus

$$|\Psi_0\rangle = S(0,-\infty)|\Phi_0\rangle. \tag{2.49}$$

By assuming that the system feels no interaction at $t \to +\infty$, the wavefunction is also non-interacting ground state $|\Phi_0\rangle$ except for a phase factor e^{iL}. Thus

$$e^{iL}|\Phi_0\rangle = S(+\infty,0)|\Psi_0\rangle = S(+\infty,-\infty)|\Phi_0\rangle. \tag{2.50}$$

From this relation, we can see

$$\begin{aligned}\langle \Psi_0| &= \langle \Phi_0|S(-\infty,0) = \langle \Phi_0|S(-\infty,+\infty)S(+\infty,0) \\ &= e^{-iL}\langle \Phi_0|S(+\infty,0).\end{aligned} \tag{2.51}$$

The field operators $\hat{\psi}_{\mathrm{H}}(x,t)$ and $\hat{\psi}_{\mathrm{H}}^{\dagger}(x,t)$ can be also converted into ones in the interaction representation as

$$\begin{aligned}\hat{\psi}_{\mathrm{H}}(x,t) &= e^{iHt}e^{-iH_0t}e^{iH_0t}\hat{\psi}(x,0)e^{-iH_0t}e^{iH_0t}e^{-iHt} \\ &= e^{iHt}e^{-iH_0t}\hat{\psi}(x,t)e^{iH_0t}e^{-iHt} \\ &= S(0,t)\hat{\psi}(x,t)S(t,0),\end{aligned} \tag{2.52}$$

and

$$\hat{\psi}_{\mathrm{H}}^{\dagger}(x,t) = S(0,t)\hat{\psi}^{\dagger}(x,t)S(t,0). \tag{2.53}$$

Substituting Eq (2.49), (2.51), (2.52), and (2.53) into Eq (2.28), the Green's function can be represented when $t > t'$ as

$$\begin{aligned}&G(x,t;x',t') \\ &= -i\frac{\langle \Phi_0|S(+\infty,0)S(0,t)\hat{\psi}(x,t)S(t,0)S(0,t')\hat{\psi}^{\dagger}(x',t')S(t',0)S(0,-\infty)|\Phi_0\rangle}{\langle \Phi_0|S(+\infty,-\infty)|\Phi_0\rangle} \\ &= -i\frac{\langle \Phi_0|S(+\infty,t)\hat{\psi}(x,t)S(t,t')\hat{\psi}^{\dagger}(x',t')S(t',-\infty)|\Phi_0\rangle}{\langle \Phi_0|S(+\infty,-\infty)|\Phi_0\rangle},\end{aligned} \tag{2.54}$$

and similarly in the case of $t' \geq t$ it is written as

$$\begin{aligned}&G(x,t;x',t') \\ &= i\frac{\langle \Phi_0|S(+\infty,0)S(0,t')\hat{\psi}^{\dagger}(x',t')S(t',0)S(0,t)\hat{\psi}(x,t)S(t,0)S(0,-\infty)|\Phi_0\rangle}{\langle \Phi_0|S(+\infty,-\infty)|\Phi_0\rangle} \\ &= i\frac{\langle \Phi_0|S(+\infty,t)\hat{\psi}^{\dagger}(x',t')S(t',t)\hat{\psi}(x,t)S(t,-\infty)|\Phi_0\rangle}{\langle \Phi_0|S(+\infty,-\infty)|\Phi_0\rangle}.\end{aligned} \tag{2.55}$$

Combining Eq (2.54) and (2.55), we obtain the following relation regardless of the time-ordering

$$iG(x,t;x',t') = \frac{\langle \Phi_0 | T\{S(+\infty,-\infty)\hat{\psi}(x,t)\hat{\psi}^{\dagger}(x',t')\} | \Phi_0 \rangle}{\langle \Phi_0 | S(+\infty,-\infty) | \Phi_0 \rangle}. \tag{2.56}$$

Here the factor i is put on the left-hand side to simplify the complicated terms in the right-hand side.

In Eq (2.56), all of the effects of interaction $\hat{V}$ are gathered in $S(+\infty,-\infty)$. According to Eq (2.46), it can be represented as follows

$$\begin{aligned} S(+\infty,-\infty) &= T\exp\left(-i\int_{-\infty}^{+\infty} \mathrm{d}t_1 \hat{V}(t_1)\right) \\ &= 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{+\infty} \mathrm{d}t_1 \int_{-\infty}^{+\infty} \mathrm{d}t_2 \cdots \int_{-\infty}^{+\infty} \mathrm{d}t_n T\{\hat{V}(t_1)\hat{V}(t_2)\cdots\hat{V}(t_n)\}. \end{aligned} \tag{2.57}$$

Now, once the interaction $\hat{V}$ is selected, the fully interacting Green's function can be obtained in chosen order of expansion n in Eq (2.57).

2-2-3. Wick's theorem

As example of pertubative expansion, we will see the first order expansion of electron-electron (e-e) interaction. In second quantization, the interaction $\hat{V}_{e-e}$ is written as

$$\hat{V}_{e-e} = \frac{1}{2}\sum_{\mathbf{q}} v_{\mathbf{q}} \sum_{\mathbf{k},\mathbf{k}',\sigma,\sigma'} c^{\dagger}_{\mathbf{k}+\mathbf{q}\sigma 1} c^{\dagger}_{\mathbf{k}'-\mathbf{q}\sigma 2} c_{\mathbf{k}'\sigma 2} c_{\mathbf{k}\sigma 1}, \tag{2.58}$$

where $v_{\mathbf{q}}$ is Fourier transform of Coulomb interaction $v_{\mathbf{q}} = \frac{4\pi}{|\mathbf{q}|^2}$. Thus the first-expansion term of Eq (2.57), $S^{(1)}(+\infty,-\infty)$ is written as

$$S^{(1)}(+\infty,-\infty)$$
$$= -\frac{i}{2}\int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{q}} v_{\mathbf{q}} \sum_{\mathbf{k},\mathbf{k}',\sigma,\sigma'} c^{\dagger}_{\mathbf{k}+\mathbf{q}\sigma 1}(t_1) c^{\dagger}_{\mathbf{k}'-\mathbf{q}\sigma 2}(t_1) c_{\mathbf{k}'\sigma 2}(t_1) c_{\mathbf{k}\sigma 1}(t_1). \tag{2.59}$$

After substituting Eq (2.23) and Eq (2.59) into Eq (2.56), the diagonal term with $\lambda = (\mathbf{p}, \sigma)$ of the numerator in the right-hand side is written as

$$\langle \Phi_0 | T\{S^{(1)}(+\infty,-\infty) c_{\mathbf{p}\sigma}(t) c^{\dagger}_{\mathbf{p}\sigma}(t')\} | \Phi_0 \rangle$$
$$= -\frac{i}{2}\int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{q}} \sum_{\mathbf{k},\mathbf{k}',\sigma 1,\sigma 2} v_{\mathbf{q}} \times \tag{2.60}$$
$$\left\langle \Phi_0 \middle| T\left\{ c^{\dagger}_{\mathbf{k}+\mathbf{q}\sigma 1}(t_1) c^{\dagger}_{\mathbf{k}'-\mathbf{q}\sigma 2}(t_1) c_{\mathbf{k}'\sigma 2}(t_1) c_{\mathbf{k}\sigma 1}(t_1) c_{\mathbf{p}\sigma}(t) c^{\dagger}_{\mathbf{p}\sigma}(t') \right\} \middle| \Phi_0 \right\rangle,$$

thus now what we have to calculate is the *T*-product of six Fermion operators. In the following we provide the treatment for solving such *T*-product.

Expectation value of collection of creation and annihilation operators is nonzero only when all the states created by $c_n^{\dagger}$ are destructed by annihilation operators c_m. Thus, the number of creation operators must be the same as that of annihilation operators, and each creation operator must be paired to a corresponding annihilation operator. When *N* creation and *N* annihilation operators exist, the number of possible parings is *N*!. For example, $\langle \Phi_0 | T\{c_\alpha(t_1) c^{\dagger}_\beta(t_1) c_\gamma(t) c^{\dagger}_\delta(t')\} | \Phi_0 \rangle$ has two creation and two annihilation operators, thus 2!=2 parings are possible, i.e., ($\alpha = \beta,\ \gamma = \delta$) or ($\alpha = \delta,\ \gamma = \beta$).

T-product of more than two creation and annihilation operators can be simplified based on the "Wick's theorem". This theorem states that *T*-product can be decomposed into more

than one term consisting of multiples of time-ordering pairs such as $\langle \Phi_0 | T\{c_{\mathbf{p}\sigma}(t) c^{\dagger}_{\mathbf{p}\sigma}(t')\} | \Phi_0 \rangle$ or $\langle \Phi_0 | T\{c^{\dagger}_{\mathbf{p}\sigma}(t) c_{\mathbf{p}\sigma}(t)\} | \Phi_0 \rangle$. The number of terms is the number of possible pairings *N*!. Signs of the terms depend on how many interchanges of neighboring Fermion operators are required to obtain the aimed ordering: an odd number of interchanges gives a negative sign. This theorem is proven using the anticommutator rule and also using the fact that the expectation value is zero in normal ordering of creation and annihilation operators.[18]

In decomposing *T*-product by Wick's theorem, we have to take account of the order of $c_n(t)$ and $c_n^{\dagger}(t')$ in time-ordering pairs.

i) $t \neq t'$

$c_n(t)$ is on the left side of $c_n^{\dagger}(t')$ to be consistent with the definition of non-interacting Green's function $iG_0^n(t - t_1) = \langle \Phi_0 | T\{c_n(t) c_n^{\dagger}(t')\} | \Phi_0 \rangle$ as shown in Eq (2.32).

ii) $t = t'$

$c_n(t) c_n^{\dagger}(t)$ becomes a time-independent operator $c_n c_n^{\dagger}$. In this case, $c_n(t)$ is on the right side of $c_n^{\dagger}(t')$ to be consistent with Fermi-Dirac distribution function $f_n = \langle \Phi_0 | c_n^{\dagger} c_n | \Phi_0 \rangle$ (see Eq (2.36)).

For example, $\left\langle \Phi_0 \middle| T\{c_\alpha(t_1) c_\beta^{\dagger}(t_1) c_\gamma(t) c_\delta^{\dagger}(t')\} \middle| \Phi_0 \right\rangle$ is decomposed into two terms consisting of multiples of two time-ordering pairs as

$$\begin{aligned}
&\left\langle \Phi_0 \middle| T\{c_\alpha(t_1)c_\beta^\dagger(t_1)c_\gamma(t)c_\delta^\dagger(t')\} \middle| \Phi_0 \right\rangle \\
&\quad = -\delta_{\alpha,\beta}\delta_{\gamma,\delta}\langle \Phi_0 | T\{c_\alpha^\dagger(t_1)c_\alpha(t_1)\} | \Phi_0 \rangle \langle \Phi_0 | T\{c_\gamma(t)c_\gamma^\dagger(t')\} | \Phi_0 \rangle \\
&\qquad -\delta_{\alpha,\delta}\delta_{\beta,\gamma}\langle \Phi_0 | T\{c_\alpha(t_1)c_\alpha^\dagger(t')\} | \Phi_0 \rangle \langle \Phi_0 | T\{c_\gamma(t)c_\gamma^\dagger(t_1)\} | \Phi_0 \rangle \\
&\quad = -\delta_{\alpha,\beta}\delta_{\gamma,\delta}[f_a iG_0^\gamma(t-t')] - \delta_{\alpha,\delta}\delta_{\beta,\gamma}[iG_0^\alpha(t_1-t')iG_0^\gamma(t-t_1)]
\end{aligned}$$

where the negative sign in the first term derives from one interchange to gain $c_\beta^\dagger(t_1)c_\alpha(t_1)c_\gamma(t)c_\delta^\dagger(t')$ from $c_\alpha(t_1)c_\beta^\dagger(t_1)c_\gamma(t)c_\delta^\dagger(t')$. Similarly, the negative sign in the second term derives from three interchanges to get $c_\alpha(t_1)c_\delta^\dagger(t')c_\gamma(t)c_\beta^\dagger(t_1)$.

In the case of the *T*-product in Eq (2.60), it is decomposed into 3! = 6 terms consisting of three time-ordering pairs. The indexes of momentum ($\mathbf{k}, \mathbf{k}', \mathbf{q}, \mathbf{p}$) and spin ($\sigma_1, \sigma_2, \sigma$) should be determined to make pairs of creation and annihilation operators. Following the same procedure as the previous example, it can be decomposed as,

$$\begin{aligned}
&\left\langle \Phi_0 \middle| T\left\{c_{\mathbf{k}+\mathbf{q}\sigma 1}^\dagger(t_1)c_{\mathbf{k}'-\mathbf{q}\sigma 2}^\dagger(t_1)c_{\mathbf{k}'\sigma 2}(t_1)c_{\mathbf{k}\sigma 1}(t_1)c_{\mathbf{p}\sigma}(t)c_{\mathbf{p}\sigma}^\dagger(t')\right\} \middle| \Phi_0 \right\rangle \\
&= \left\langle \Phi_0 \middle| T\left\{\underset{\text{a}}{c_{\mathbf{k}+\mathbf{q}\sigma 1}^\dagger(t_1)}\underset{\text{a'}}{c_{\mathbf{k}'\sigma 2}(t_1)}\underset{\text{b}}{c_{\mathbf{k}'-\mathbf{q}\sigma 2}^\dagger(t_1)}\underset{\text{b'}}{c_{\mathbf{k}\sigma 1}(t_1)}\underset{\text{c}}{c_{\mathbf{p}\sigma}^\dagger(t')}\underset{\text{c'}}{c_{\mathbf{p}\sigma}(t)}\right\} \middle| \Phi_0 \right\rangle \\
&= +\delta_{\mathbf{k},\mathbf{p}}\delta_{\mathbf{q},0}\delta_{\sigma 1,\sigma}[f_{\mathbf{k}',\sigma 2}iG_0^{\mathbf{p}\sigma}(t-t_1)iG_0^{\mathbf{p}\sigma}(t_1-t')] \\
&\quad -\delta_{\mathbf{k},\mathbf{p}-\mathbf{q}}\delta_{\mathbf{k}',\mathbf{p}}\delta_{\sigma 1,\sigma}\delta_{\sigma 2,\sigma}[f_{\mathbf{p}-\mathbf{q},\sigma}iG_0^{\mathbf{p}\sigma}(t-t_1)iG_0^{\mathbf{p}\sigma}(t_1-t')] \\
&\quad +\delta_{\mathbf{q},0}[f_{\mathbf{k},\sigma 1}f_{\mathbf{k}',\sigma 2}iG_0^{\mathbf{p}\sigma}(t-t')] \\
&\quad +\delta_{\mathbf{k}',\mathbf{p}}\delta_{\mathbf{q},0}\delta_{\sigma 2,\sigma}[f_{\mathbf{k},\sigma 1}iG_0^{\mathbf{p}\sigma}(t-t_1)iG_0^{\mathbf{p}\sigma}(t_1-t')] \\
&\quad -\delta_{\mathbf{k},\mathbf{p}}\delta_{\mathbf{k}',\mathbf{p}+\mathbf{q}}\delta_{\sigma 1,\sigma}\delta_{\sigma 2,\sigma}[f_{\mathbf{p}+\mathbf{q},\sigma}iG_0^{\mathbf{p}\sigma}(t-t_1)iG_0^{\mathbf{p}\sigma}(t_1-t')] \\
&\quad -\delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}}\delta_{\sigma 1,\sigma 2}[f_{\mathbf{k},\sigma 1}f_{\mathbf{k}+\mathbf{q},\sigma 1}iG_0^{\mathbf{p}\sigma}(t-t')].
\end{aligned} \tag{2.61}$$

Once we assign indexes of a, b, and c for creation operators and a', b', and c' for annihilation operators, as shown in Eq (2.61), the pairing of these six terms are represented as (a=c', b=a', c=b'), (a=c', b=b', c=a'), (a=b', b=a', c=c'), (a=b', b=c', c=a'), (a=a', b=c', c=b'), and (a=a', b=b', c=c') in the order of Eq (2.61). Substituting Eq (2.61) into Eq (2.60), finally we obtain six first-order perturbation terms, from $iG^{\mathbf{p}\sigma\,(1\mathrm{A})}$ to $iG^{\mathbf{p}\sigma\,(1\mathrm{F})}$.

$iG^{\mathbf{p}\sigma\,(1\mathrm{A})}$ and $iG^{\mathbf{p}\sigma\,(1\mathrm{D})}$ are identical and their summation represents the "Hartree" term as

$$\begin{aligned} iG^{\mathbf{p}\sigma}_{(\mathrm{Hartree})}(t-t') &= iG^{\mathbf{p}\sigma\,(1\mathrm{A})}(t-t') + iG^{\mathbf{p}\sigma\,(1\mathrm{D})}(t-t') \\ &= -i\int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{k},\sigma 1} \left[v_{\mathbf{q}=0} f_{\mathbf{k},\sigma 1} iG_0^{\mathbf{p}\sigma}(t-t_1) iG_0^{\mathbf{p}\sigma}(t_1-t') \right], \end{aligned} \tag{2.62}$$

and similarly, $iG^{\mathbf{p}\sigma\,(1\mathrm{B})} + iG^{\mathbf{p}\sigma\,(1\mathrm{E})}$ represents the "Fock" term as

$$\begin{aligned} iG^{\mathbf{p}\sigma}_{(\mathrm{Fock})}(t-t') &= iG^{\mathbf{p}\sigma\,(1\mathrm{B})}(t-t') + iG^{\mathbf{p}\sigma\,(1\mathrm{E})}(t-t') \\ &= i\int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{q}} \left[v_{\mathbf{q}} f_{\mathbf{p}-\mathbf{q},\sigma} iG_0^{\mathbf{p}\sigma}(t-t_1) iG_0^{\mathbf{p}\sigma}(t_1-t') \right]. \end{aligned} \tag{2.63}$$

$iG^{\mathbf{p}\sigma\,(1\mathrm{C})}$ and $iG^{\mathbf{p}\sigma\,(1\mathrm{F})}$ are so called *disconnected* terms including $iG_0^{\mathbf{p}\sigma}(t-t')$ as

$$iG^{\mathbf{p}\sigma\,(1\mathrm{C})}(t-t') = -\frac{i}{2}\int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{k},\mathbf{k}',\sigma 1,\sigma 2} \left[v_{\mathbf{q}=0} f_{\mathbf{k},\sigma 1} f_{\mathbf{k}',\sigma 2} iG_0^{\mathbf{p}\sigma}(t-t') \right] \tag{2.64}$$

$$iG^{\mathbf{p}\sigma\,(1\mathrm{F})}(t-t') = \frac{i}{2}\int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{q}} \sum_{\mathbf{k},\sigma 1} \left[v_{\mathbf{q}} f_{\mathbf{k},\sigma 1} f_{\mathbf{k}+\mathbf{q},\sigma 1} iG_0^{\mathbf{p}\sigma}(t-t') \right]. \tag{2.65}$$

Feynman introduced the idea of a pictorial representation of these perturbation terms. These drawings, called *Feynman diagrams*, are extremely useful for providing an insight into the physical process which these terms represent. These diagrams can be drawn using connection of the ingredients iG_0, f, and $v_{\mathbf{q}}$. G_0 can be drawn in both time-space (**k**, t) and energy-space (**k**, ω).

The main rules for constructing such diagrams are:

1) Each ingredient is assigned to the drawing as below.

2) Time is advanced from left to right: lines of $v_{\mathbf{q}}$ and f are drawn vertically since they are time-independent.

3) A factor of -1 is added for each loop of solid lines (Fermion loop). 

4) The momentum and spin (and also frequency in (**k**, ω)-space) of entering and leaving lines must be consistent at each vertex.

5) Each intermediate variables must be integrated as $\Sigma_{\mathbf{k}}$, Σ_{σ}, and $\int_{-\infty}^{+\infty} \mathrm{d}t$ (in (**k**, t)-space) or $\int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi}$ (in (**k**, ω)-space).

The resulting diagrams of six perturbation terms are shown in Fig 2-1. These diagrams are consistent with the six terms in Eq (2.62), (2.63), (2.64), and (2.65).

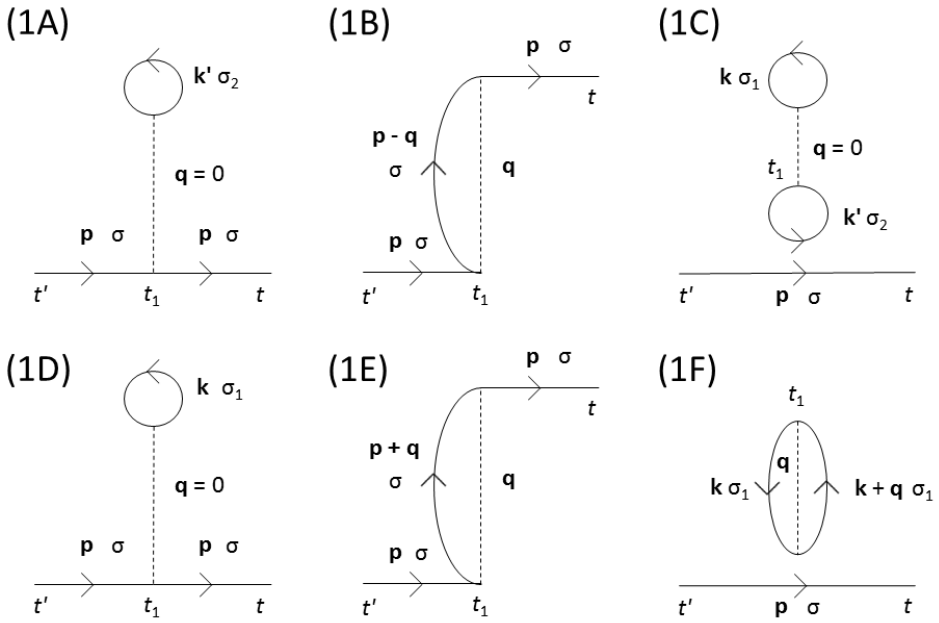


Figure 2-1. Feynman diagrams of first-order perturbation terms from e-e interaction.

2-2-4. Dyson's equation

Considering higher-order expansion, the number of terms to be considered rapidly increases, e.g., second-order terms 5!=120, and third-order terms 7!=5040. However, not every term is physically meaningful thus their reduction is possible according to the following procedure.

First step of simplification is to neglect disconnected terms. The denominator of Eq (2.56) can also be expanded by Wick's theorem. In the first-order, the denominator has two creation and two annihilation operators, thus it can be decomposed 2!=2 terms as below:

$$\begin{aligned}&\left\langle \Phi_0 \middle| T\left\{ c^{\dagger}_{\mathbf{k+q}\sigma 1}(t_1) c^{\dagger}_{\mathbf{k'-q}\sigma 2}(t_1) c_{\mathbf{k'}\sigma 2}(t_1) c_{\mathbf{k}\sigma 1}(t_1) \right\} \middle| \Phi_0 \right\rangle \\ &= \delta_{\mathbf{q},0}\left[f_{\mathbf{k},\sigma 1} f_{\mathbf{k'},\sigma 2} \right] - \delta_{\mathbf{k'},\mathbf{k+q}} \delta_{\sigma 1,\sigma 2}\left[f_{\mathbf{k},\sigma 1} f_{\mathbf{k+q},\sigma 1} \right].\end{aligned} \tag{2.66}$$

Substituting Eq (2.66) into Eq (2.56), the denominator is represented as

$$\begin{aligned}
&\langle \Phi_0 | S^{(1)}(+\infty, -\infty) | \Phi_0 \rangle \\
&= -\frac{i}{2} \int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{q}} \sum_{\mathbf{k},\mathbf{k}',\sigma 1,\sigma 2} v_{\mathbf{q}} \left\langle \Phi_0 \middle| T\left\{ c^{\dagger}_{\mathbf{k}+\mathbf{q}\sigma 1}(t_1) c^{\dagger}_{\mathbf{k}'-\mathbf{q}\sigma 2}(t_1) c_{\mathbf{k}'\sigma 2}(t_1) c_{\mathbf{k}\sigma 1}(t_1) \right\} \middle| \Phi_0 \right\rangle \\
&= -\frac{i}{2} \int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{k},\mathbf{k}',\sigma 1,\sigma 2} \left[v_{\mathbf{q}=0} f_{\mathbf{k},\sigma 1} f_{\mathbf{k}',\sigma 2} \right] + \frac{i}{2} \int_{-\infty}^{+\infty} \mathrm{d}t_1 \sum_{\mathbf{q}} \sum_{\mathbf{k},\sigma 1} v_{\mathbf{q}} \left[f_{\mathbf{k},\sigma 1} f_{\mathbf{k}+\mathbf{q},\sigma 1} \right].
\end{aligned} \tag{2.67}$$

The terms in the series for $\langle \Phi_0 | S(+\infty, -\infty) | \Phi_0 \rangle$ are called *vacuum polarization terms*. Comparing Eq (2.67) with disconnected terms in Eq (2.64) and Eq (2.65), we notice that $iG^{\mathbf{p}\sigma\,(1\mathrm{C})}$ and $iG^{\mathbf{p}\sigma\,(1\mathrm{F})}$ can be represented using the vacuum polarization terms as

$$iG^{\mathbf{p}\sigma\,(1\mathrm{C})}(t-t') + iG^{\mathbf{p}\sigma\,(1\mathrm{F})}(t-t') = iG_0^{\mathbf{p}\sigma}(t-t') \langle \Phi_0 | S^{(1)}(+\infty, -\infty) | \Phi_0 \rangle. \tag{2.68}$$

The first-order Green's function $iG^{(1)}$ is drawn in the Feynman diagram (Fig 2-2). According to Eq (2.68), disconnected diagrams are shown as a product of iG_0 and vacuum polarization diagrams. Once we consider higher-order perturbation, the *n*th-order disconnected terms represent as a product of the (*n*-*m*)th-order vacuum polarization terms and the *m*th-order connected diagrams ($0 \leq m \leq n-1$). Finally, vacuum polarization terms are completely canceled out in the infinite order and only connected diagrams are now available. This is known as "linked cluster theorem".

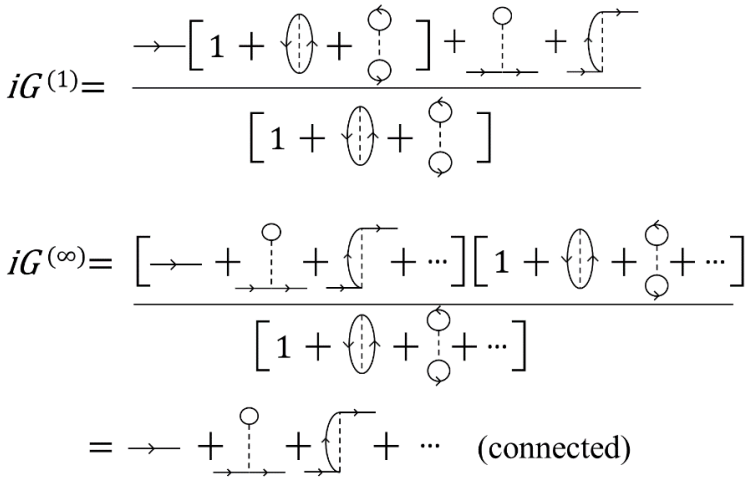


Figure 2-2. Schematic representation of linked cluster theorem

Thus Eq (2.56) is rewritten as below.

$$iG^{\mathbf{p},\sigma}(t-t') = iG_0^{\mathbf{p},\sigma}(t-t') + \sum_{n=1}^{\infty}\frac{(-i)^n}{n!}\int_{-\infty}^{+\infty}\mathrm{d}t_1\cdots\int_{-\infty}^{+\infty}\mathrm{d}t_n \times \\ \left\langle\Phi_0\middle|T\left\{\hat{V}(t_1)\cdots\hat{V}(t_n)c_{\mathbf{p}\sigma}(t)c_{\mathbf{p}\sigma}^{\dagger}(t')\right\}\middle|\Phi_0\right\rangle \text{ (connected)} \tag{2.69}$$

As shown above, 1/2 factor in Eq (2.58) is absent in both Hartree and Fock self-energies because there are two identical terms. The factor 2 derives from the fact that intermediate states $(\mathbf{k},\sigma_1)$ and $(\mathbf{k}',\sigma_2)$ are independent. Similarly, we can get rid of the 1/*n*! factor in Eq (2.57) in *n*th-order expansion because there are just *n*! terms exactly identical in the *n*th-order *T*-product. As a result, further simplification is achieved.

$$iG^{\mathbf{p},\sigma}(t-t') = iG_0^{\mathbf{p},\sigma}(t-t') + \sum_{n=1}^{\infty}(-i)^n\int_{-\infty}^{+\infty}\mathrm{d}t_1\cdots\int_{-\infty}^{+\infty}\mathrm{d}t_n \times \\ \left\langle\Phi_0\middle|T\left\{\hat{V}(t_1)\cdots\hat{V}(t_n)c_{\mathbf{p}\sigma}(t)c_{\mathbf{p}\sigma}^{\dagger}(t')\right\}\middle|\Phi_0\right\rangle, \text{(different connected)} \tag{2.70}$$

Accordingly, first-order Green's function, $iG^{(1)}$ is represented by Hartree and Fock terms;

$$iG^{\mathbf{p},\sigma(1)}(t-t') = iG_0^{\mathbf{p},\sigma}(t-t') + iG_{\mathrm{Hartree}}^{\mathbf{p}\sigma}(t-t') + iG_{\mathrm{Fock}}^{\mathbf{p}\sigma}(t-t').$$

For connected terms the *self-energy*, Σ, is defined as the remaining part from incoming and outgoing propagators;

$$iG^{\mathbf{p},\sigma}(t-t') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt_1 dt_2 \left[iG_0^{\mathbf{p},\sigma}(t-t_1)\right]\left[-i\Sigma^{\mathbf{p},\sigma}(t_1-t_2)\right]\left[iG_0^{\mathbf{p},\sigma}(t_2-t')\right].$$

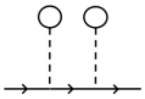
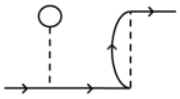
Accordingly, the self-energies of Hartree and Fock terms are defined as

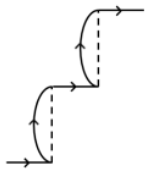
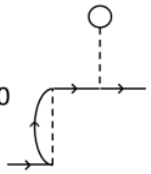
$$\Sigma_{\text{Hartree}}^{\mathbf{p},\sigma}(t=0) = \sum_{\mathbf{k},\sigma 1} v_{\mathbf{q}=0} f_{\mathbf{k}\sigma 1}, \tag{2.71}$$

$$\Sigma_{\text{Fock}}^{\mathbf{p},\sigma}(t=0) = -\sum_{\mathbf{q}} v_{\mathbf{q}} f_{\mathbf{p}-\mathbf{q}\sigma}. \tag{2.72}$$

Notice that these self-energies are time-independent because t_1 is equal to t_2 in the first-order case.

Final procedure to reduce considered perturbation terms is to avoid higher-order diagrams consisting of the lower order self-energies and zeroth-order Green's functions. For example, in the second-order terms there are diagrams as:

$G_0\Sigma_{\text{Hartree}}G_0\Sigma_{\text{Hartree}}G_0$  $G_0\Sigma_{\text{Hartree}}G_0\Sigma_{\text{Fock}}G_0$ 

$G_0\Sigma_{\text{Fock}}G_0\Sigma_{\text{Fock}}G_0$  $G_0\Sigma_{\text{Fock}}G_0\Sigma_{\text{Hartree}}G_0$ 

These terms are called *improper* or *reducible* diagrams. In the infinite-order, such improper diagrams can be factorized as

$$G = G_0 + G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0\Sigma G_0 + \cdots$$

$$= G_0 + G_0\Sigma[G_0 + G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0 + \cdots]$$

$$= G_0 + G_0\Sigma G.$$

Finally, the expansion of Green's function results in *Dyson's equation*;

$$iG^{\mathbf{p},\sigma}(t-t') = iG_0^{\mathbf{p},\sigma}(t-t') + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dt_1 dt_2 \left[iG_0^{\mathbf{p},\sigma}(t-t_1)\right][-i\Sigma^{\mathbf{p},\sigma}(t_1-t_2)][iG^{\mathbf{p},\sigma}(t_2-t')], \tag{2.73}$$

where only *proper* self-energies, which are not factorized into the lower-order terms, are considered. The schematic representation of Dyson's equation is shown in Fig. 2-3.

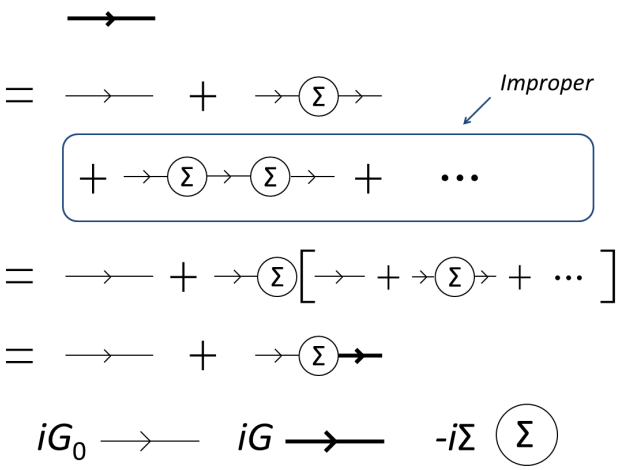


Figure 2-3. Schematic representation of Dyson's equation.

By Fourier transformation of Eq (2.73), Dyson's equation in energy-space is obtained;

$$G^{\mathbf{p},\sigma}(\omega) = G_0^{\mathbf{p},\sigma}(\omega) + G_0^{\mathbf{p},\sigma}(\omega)\Sigma^{\mathbf{p},\sigma}(\omega)G^{\mathbf{p},\sigma}(\omega). \tag{2.74}$$

From Eq (2.38), Dyson's equation can be also written as

$$G^{\mathbf{p},\sigma}(\omega) = \frac{1}{\left[G_0^{\mathbf{p},\sigma}(\omega)\right]^{-1} - \Sigma^{\mathbf{p},\sigma}(\omega)} = \frac{1}{\omega - \varepsilon_{\mathbf{p}\sigma} \pm i\delta - \Sigma^{\mathbf{p},\sigma}(\omega)}, \tag{2.75}$$

where the sign of $i\delta$ is positive (negative) for unoccupied (occupied) states. Using this equation, interacting Green's function is determined once we define the self-energy.

2-2-5. Quasi-particle approximation

Once we know the interacting Green's function, quasiparticle (QP) energy, $\varepsilon_\lambda^{\mathrm{QP}}$ is determined as a root of the right-hand side of Eq (2.75),

$$\varepsilon_\lambda^{\mathrm{QP}} - \varepsilon_\lambda^0 - \Sigma^\lambda\left(\varepsilon_\lambda^{\mathrm{QP}}\right) = 0. \tag{2.76}$$

By means of a Taylor expansion of $\Sigma^\lambda(\omega)$ around $\omega = \varepsilon_\lambda^0$ to the first order, $\Sigma^\lambda\left(\varepsilon_\lambda^{\mathrm{QP}}\right)$ is approximated as

$$\Sigma^\lambda\left(\varepsilon_\lambda^{\mathrm{QP}}\right) \approx \Sigma^\lambda\left(\varepsilon_\lambda^0\right) + \left.\frac{\partial\Sigma^\lambda}{\partial\omega}\right|_{\omega=\varepsilon_\lambda^0}\left(\varepsilon_\lambda^{\mathrm{QP}} - \varepsilon_\lambda^0\right). \tag{2.77}$$

Substituting Eq (2.77) into Eq (2.76), we finally obtain

$$\varepsilon_\lambda^{\mathrm{QP}} = \varepsilon_\lambda^0 + Z_\lambda\Sigma^\lambda\left(\varepsilon_\lambda^0\right), \tag{2.78}$$

where $Z_\lambda = \left(1 - \left.\frac{\partial\Sigma^\lambda}{\partial\omega}\right|_{\omega=\varepsilon_\lambda^0}\right)^{-1}$ is the renormalization factor. This procedure is called quasiparticle approximation (QPA).

For example, using Eq (2.71) and (2.72), QP energy in the first-order expansion is represented as

$$\begin{aligned}\varepsilon_{\mathbf{p},\sigma}^{\mathrm{HF}} &= \varepsilon_{\mathbf{p},\sigma}^0 + \Sigma_{\mathrm{Hartree}}^{\mathbf{p},\sigma} + \Sigma_{\mathrm{Fock}}^{\mathbf{p},\sigma} \\ &= \varepsilon_{\mathbf{p},\sigma}^0 + \sum_{\mathbf{k},\sigma 1} v_{\mathbf{q}=0} f_{\mathbf{k}\sigma 1} - \sum_{\mathbf{q}} v_{\mathbf{q}} f_{\mathbf{p}-\mathbf{q}\sigma}.\end{aligned} \tag{2.79}$$

This single-particle energy is the solution of Hartree-Fock approximation (HFA). In HFA the self-energies are static, thus $Z_{\mathbf{p},\sigma} = 1$. Dynamical correlation interaction, which is required for accurate calculation of bandgap, is considered in the higher-order expansions in the next section.

2-3. GW approximation

The first-order perturbation, i.e., HFA, is not satisfactory to calculate bandgaps since the self-energy includes no electron-electron correlation effect. In this section, GW approximation, a powerful approach to calculate bandgaps, is explained. In the GW approximation dynamical correlation interaction is considered by partial summation of "ring"-shaped self-energy diagrams up to the infinite order.

2-3-1. Self-energy in GW approximation

The lowest self-energy of the ring diagram, which is one of the second order terms, is shown in Fig. 2-4. This self-energy, $\Sigma^{\mathbf{p}}_{\mathrm{ring}}(\omega)$ is formulated as

$$-i\Sigma^{\mathbf{p}}_{\mathrm{ring}}(\omega) = \sum_{\mathbf{q}} \int \frac{d\omega'}{2\pi} \left(-iv_{\mathbf{q}}\right)^2 \left[iG_0^{\mathbf{p}-\mathbf{q}}(\omega-\omega')\right]\left[iP_0^{\mathbf{q}}(\omega')\right], \qquad (2.80)$$

where $P_0^{\mathbf{q}}(\omega')$ is the ring part of the diagram in Fig 2-4 representing *polarizability* of free electron and defined as

$$\begin{aligned} iP_0^{\mathbf{q}}(\omega') &= (-1)\times 2\sum_{\mathbf{q}'} \int \frac{d\varepsilon}{2\pi}\left[iG_0^{\mathbf{q}+\mathbf{q}'}(\omega'+\varepsilon)\right]\left[iG_0^{\mathbf{q}'}(\varepsilon)\right] \\ &= 2i\sum_{\mathbf{q}'}\left[\frac{f_{\mathbf{q}'}(1-f_{\mathbf{q}+\mathbf{q}'})}{\omega'-\varepsilon_{\mathbf{q}+\mathbf{q}'}+\varepsilon_{\mathbf{q}'}+i\delta} - \frac{f_{\mathbf{q}+\mathbf{q}'}(1-f_{\mathbf{q}'})}{\omega'-\varepsilon_{\mathbf{q}+\mathbf{q}'}+\varepsilon_{\mathbf{q}'}-i\delta}\right], \end{aligned} \qquad (2.81)$$

where the factor -1 comes from the Fermion loop and the factor 2 is from up and down spins. The final representation is derived by substituting Eq (2.38) and residue theorem for integration of ε.[17]

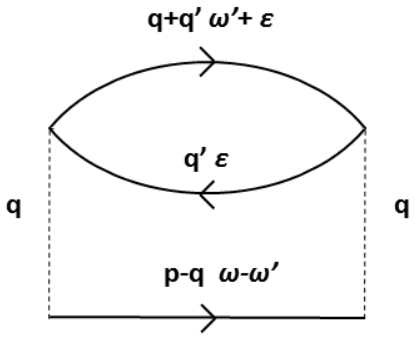


Figure 2-4. Lowest self-energy of the ring diagram.

The polarizability introduces dynamical screening effects to the bare electron-electron interaction due to the surrounding electrons. This self-energy is, however, divergent at $\mathbf{q} \to 0$ due to the denominator of $v_{\mathbf{q}} = \frac{4\pi}{|\mathbf{q}|^2}$. Such issue is fixed via the partial summation of the ring diagrams up to the infinite order.[17]

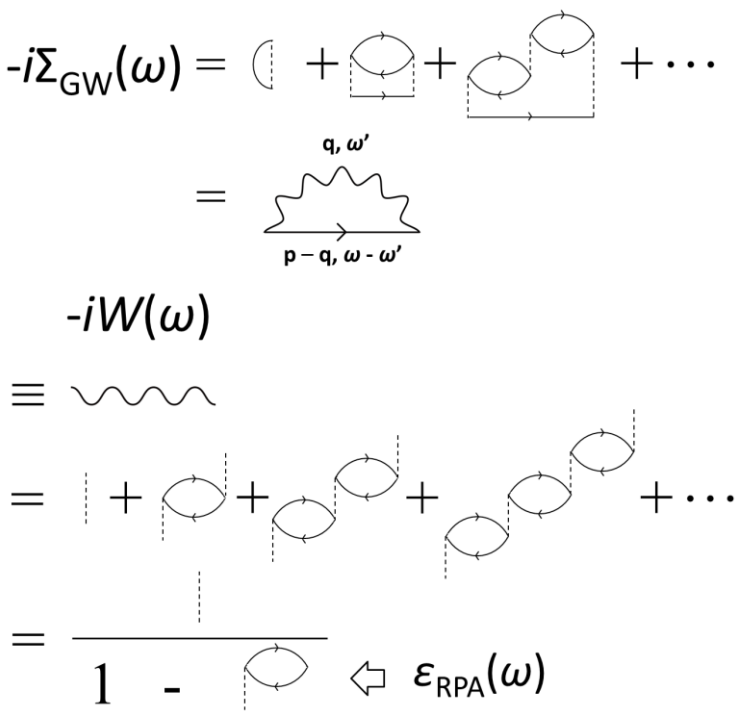


Figure 2-5. Self-energy of GW approximation and screened interaction

As shown in Fig. 2-5, the partial summation of ring diagrams with Fock term results in self-energy of GW approximation written as

$$-i\Sigma_{\mathrm{GW}}^{\mathbf{p}}(\omega) = \frac{1}{2\pi}\sum_{\mathbf{q}}\int d\omega' \left[iG_0^{\mathbf{p}-\mathbf{q}}(\omega-\omega')\right]\left[-iW_{\mathbf{q}}(\omega')\right], \tag{2.82}$$

where the wiggle, $W_{\mathbf{q}}(\omega')$ is the screened Coulomb interaction in the "random phase approximation" (RPA) defined as

$$-iW_{\mathbf{q}}(\omega') = \frac{-iv_{\mathbf{q}}}{\varepsilon_{\mathrm{RPA}}^{\mathbf{q}}(\omega')}. \tag{2.83}$$

$\varepsilon_{\mathrm{RPA}}^{\mathbf{q}}(\omega')$ is the dielectric function in RPA;

$$\varepsilon_{\mathrm{RPA}}^{\mathbf{q}}(\omega') = 1 - v_{\mathbf{q}}P_0^{\mathbf{q}}(\omega'). \tag{2.84}$$

The "GW" approximation assumes its name by this formula as product of "*G*" and "*W*".

GW approximation is also derived as the simplest approximation to the Hedin's equation (Schwinger's functional derivative method).[19] Hedin's equations relate five quantities: polarizability P, screening interaction W, self-energy Σ, Green's function G, and vertex function Γ;

$$G(12) = G_0(12) + \int G_0(13)\Sigma(34)G(42)\,d(34), \tag{2.85}$$

$$\Gamma(123) = \delta(12)\delta(13) + \int \frac{\delta\Sigma(12)}{\delta G(45)}G(46)G(75)\Gamma(673)d(4567), \tag{2.86}$$

$$P(12) = -i\int G(13)G(41)\,\Gamma(342)d(34), \tag{2.87}$$

$$W(12) = v(12) + \int v(13)P(34)\,W(42)d(34), \tag{2.88}$$

$$\Sigma(12) = i\int G(13)\Gamma(324)W(41)\,d(34), \tag{2.89}$$

where space, time, and spin coordinates are indicated in a concise form with numbers, i.e., $1 = (\mathbf{r}_1, t_1, \sigma_1)$. Eq (2.85) is the Dyson's equation itself. Derivation of the Hedin's equations is, for instance, found in Ref [19].

The simplest approximation on these five equations is to assume the second term in the right-hand side of Eq (2.86) is zero:

$$\Gamma(123) = \delta(12)\delta(13), \tag{2.90}$$

which corresponds to neglect the vertex corrections. Accordingly, Eqs (2.87) and (2.89) are extremely simplified as

$$P(12) = -iG(12)G(21), \tag{2.91}$$

$$\Sigma(12) = iG(12)W(21). \tag{2.92}$$

After the Fourier transform from *t*-space to ω-space, these representations are exactly the same as Eqs (2.81) and (2.82). The screened interaction in Eq (2.88) also turns to be the same form as Eqs (2.83) and (2.84).

2-3-2. Practical calculation

In practice, the zeroth-order Green's function is constituted by Bloch states computed at the KS-DFT level with energies $\{\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}\}$ and wavefunctions $\{\psi_{n\mathbf{k}}^{\mathrm{KS}}(\mathbf{r})\}$;

$$G_0(\mathbf{r},t;\mathbf{r}',t') = -i\sum_{n,\mathbf{k}} \psi_{n\mathbf{k}}^{\mathrm{KS}}(\mathbf{r})\psi_{n\mathbf{k}}^{\mathrm{KS}*}(\mathbf{r}')\left\langle \Phi_{\mathrm{KS}}\middle| T\left\{c_{n\mathbf{k}}(t)c_{n\mathbf{k}}^{\dagger}(t')\right\}\middle|\Phi_{\mathrm{KS}}\right\rangle$$

After the Fourier-transform, it is represented in the ω-space as

$$\begin{aligned} G_0(\mathbf{r},\mathbf{r}',\omega) &= \sum_{n,\mathbf{k}} \psi_{n\mathbf{k}}^{\mathrm{KS}}(\mathbf{r})\psi_{n\mathbf{k}}^{\mathrm{KS}*}(\mathbf{r}')\left[\frac{1-f_{n\mathbf{k}}}{\omega-\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}+i\delta}+\frac{f_{n\mathbf{k}}}{\omega-\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}-i\delta}\right] \\ &= \sum_{n,\mathbf{k}} \frac{\psi_{n\mathbf{k}}^{\mathrm{KS}}(\mathbf{r})\psi_{n\mathbf{k}}^{\mathrm{KS}*}(\mathbf{r}')}{\omega-\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}+i\delta\mathrm{sgn}[\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}-\mu]}, \end{aligned} \tag{2.93}$$

where μ is the Fermi energy.[20]

Polarization in RPA, $P_0(\mathbf{r},\mathbf{r}',\omega)$ is obtained from Eqs (2.81) and (2.93) as follows;

$$\begin{aligned} &P_0(\mathbf{r},\mathbf{r}',\omega) \\ &= -2i\int \frac{d\varepsilon}{2\pi} G_0(\mathbf{r},\mathbf{r}',\omega+\varepsilon)G_0(\mathbf{r}',\mathbf{r},\varepsilon) \\ &= 2\sum_{n\mathbf{k}}\sum_{n'\mathbf{k}'} \psi_{n\mathbf{k}}^{\mathrm{KS}}(\mathbf{r})\psi_{n\mathbf{k}}^{\mathrm{KS}*}(\mathbf{r}')\psi_{n'\mathbf{k}'}^{\mathrm{KS}}(\mathbf{r}')\psi_{n'\mathbf{k}'}^{\mathrm{KS}*}(\mathbf{r})\times \\ &\qquad \left[\frac{f_{n'\mathbf{k}'}(1-f_{n\mathbf{k}})}{\omega-\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}+\varepsilon_{n'\mathbf{k}'}^{\mathrm{KS}}+i\delta}-\frac{f_{n\mathbf{k}}(1-f_{n'\mathbf{k}'})}{\omega-\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}+\varepsilon_{n'\mathbf{k}'}^{\mathrm{KS}}-i\delta}\right] \\ &= 2\sum_{n\mathbf{k}}^{\mathrm{unocc}}\sum_{n'\mathbf{k}'}^{\mathrm{occ}} \psi_{n\mathbf{k}}^{\mathrm{KS}}(\mathbf{r})\psi_{n\mathbf{k}}^{\mathrm{KS}*}(\mathbf{r}')\psi_{n'\mathbf{k}'}^{\mathrm{KS}}(\mathbf{r}')\psi_{n'\mathbf{k}'}^{\mathrm{KS}*}(\mathbf{r})\times \\ &\qquad \left[\frac{1}{\omega-\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}+\varepsilon_{n'\mathbf{k}'}^{\mathrm{KS}}+i\delta}-\frac{1}{\omega+\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}-\varepsilon_{n'\mathbf{k}'}^{\mathrm{KS}}-i\delta}\right]. \end{aligned} \tag{2.94}$$

Noting that $P_0(\mathbf{r}+\mathbf{T},\mathbf{r}'+\mathbf{T}) = P_0(\mathbf{r},\mathbf{r}')$, P_0 can be expanded in Bloch basis functions as[19]

$$P_0(\mathbf{r},\mathbf{r}',\omega)=\frac{1}{\Omega}\sum_{\mathbf{q},\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}e^{-i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}P_0^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega), \tag{2.95}$$

where Ω is the unit cell volume. Here $P_0^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$ is represented as:

$$\begin{aligned}P_0^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)&=\iint d\mathbf{r}d\mathbf{r}'\,e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}}e^{i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}P_0(\mathbf{r},\mathbf{r}',\omega)\\ &=2\sum_{\mathbf{k}}\sum_{n}^{\text{unocc}}\sum_{n'}^{\text{occ}}\left\langle\psi^{\text{KS}}_{n'\mathbf{k}-\mathbf{q}}\middle|e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}}\middle|\psi^{\text{KS}}_{n\mathbf{k}}\right\rangle\left\langle\psi^{\text{KS}}_{n\mathbf{k}}\middle|e^{i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}\middle|\psi^{\text{KS}}_{n'\mathbf{k}-\mathbf{q}}\right\rangle\\ &\quad\times\left[\frac{1}{\omega-\varepsilon^{\text{KS}}_{n\mathbf{k}}+\varepsilon^{\text{KS}}_{n'\mathbf{k}-\mathbf{q}}+i\delta}-\frac{1}{\omega+\varepsilon^{\text{KS}}_{n\mathbf{k}}-\varepsilon^{\text{KS}}_{n'\mathbf{k}-\mathbf{q}}-i\delta}\right].\end{aligned} \tag{2.96}$$

In real space, the dielectric function $\varepsilon(\mathbf{r},\mathbf{r}',\omega)$ is defined as

$$\varepsilon(\mathbf{r},\mathbf{r}',\omega)=\delta(\mathbf{r}-\mathbf{r}')-\int d\mathbf{r}''v(\mathbf{r}-\mathbf{r}'')P(\mathbf{r}'',\mathbf{r}'), \tag{2.97}$$

and it can be also expanded in Bloch basis functions as

$$\begin{aligned}\varepsilon^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)&=\iint d\mathbf{r}d\mathbf{r}'\,e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}}e^{i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}\varepsilon(\mathbf{r},\mathbf{r}',\omega)\\ &=\delta_{\mathbf{G},\mathbf{G}'}-\frac{4\pi}{|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|}P_0^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega),\end{aligned} \tag{2.98}$$

which can be calculated by substituting Eq (2.96) into $P_0^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$.

Once $\varepsilon^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$ is obtained, the screened interaction $W(\mathbf{r},\mathbf{r}',\omega)$ can be calculated through

$$W(\mathbf{r},\mathbf{r}',\omega)=\frac{1}{\Omega}\sum_{\mathbf{q},\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}e^{-i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega), \tag{2.99}$$

and

$$W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)=4\pi\frac{1}{|\mathbf{q}+\mathbf{G}|}[\varepsilon^{-1}(\mathbf{q},\omega)]^{\mathbf{G},\mathbf{G}'}\frac{1}{|\mathbf{q}+\mathbf{G}'|}, \tag{2.100}$$

where ε^{-1} is the inverse of dielectric matrix (Eq (2.98)).

Using Eqs (2.93) and (2.99), self-energy in GW approximation is represented as

$$\begin{aligned}
&\Sigma^{\mathrm{GW}}(\mathbf{r},\mathbf{r}',\omega) \\
&= \frac{i}{2\pi}\int_{-\infty}^{+\infty} d\omega'\, G_0(\mathbf{r},\mathbf{r}',\omega-\omega')W(\mathbf{r},\mathbf{r}',\omega') \\
&= \frac{i}{2\pi}\int_{-\infty}^{+\infty} d\omega' \left[\sum_{n',\mathbf{k}'} \frac{\psi^{\mathrm{KS}}_{n'\mathbf{k}'}(\mathbf{r})\psi^{\mathrm{KS}*}_{n'\mathbf{k}'}(\mathbf{r}')}{\omega-\omega'-\varepsilon^{\mathrm{KS}}_{n'\mathbf{k}'}+i\delta\mathrm{sgn}[\varepsilon^{\mathrm{KS}}_{n'\mathbf{k}'}-\mu]}\right] \\
&\qquad \times\left[\frac{1}{\Omega}\sum_{\mathbf{q},\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}e^{-i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega')\right],
\end{aligned} \tag{2.101}$$

thus $\Sigma^{\mathrm{GW}}(\mathbf{r},\mathbf{r}',\omega)$ is calculated once $W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$ is obtained. Finally, the matrix element of the state (n, $\mathbf{k}$) is obtained by the following equation

$$\begin{aligned}
&\Sigma^{\mathrm{GW}}_{n\mathbf{k}}(\omega) \\
&= \iint d\mathbf{r}d\mathbf{r}'\,\psi^{\mathrm{KS}*}_{n\mathbf{k}}(\mathbf{r})\psi^{\mathrm{KS}}_{n\mathbf{k}}(\mathbf{r}')\Sigma^{\mathrm{GW}}(\mathbf{r},\mathbf{r}',\omega) \\
&= i\sum_{n',\mathbf{q},\mathbf{G},\mathbf{G}'}\int_{-\infty}^{+\infty}\frac{d\omega'}{2\pi\Omega}W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega')\frac{\left\langle\psi^{\mathrm{KS}}_{n\mathbf{k}}\middle|e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}\middle|\psi^{\mathrm{KS}}_{n'\mathbf{k}-\mathbf{q}}\right\rangle\left\langle\psi^{\mathrm{KS}}_{n'\mathbf{k}-\mathbf{q}}\middle|e^{-i(\mathbf{q}+\mathbf{G}')\mathbf{r}'}\middle|\psi^{\mathrm{KS}}_{n\mathbf{k}}\right\rangle}{\omega-\omega'-\varepsilon^{\mathrm{KS}}_{n'\mathbf{k}-\mathbf{q}}+i\delta\mathrm{sgn}[\varepsilon^{\mathrm{KS}}_{n'\mathbf{k}-\mathbf{q}}-\mu]}.
\end{aligned} \tag{2.102}$$

One of the major computational efforts in self-energy calculations is to compute $W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega')$ on ω' grids. One of the methods to reduce the computational burden is represented by the so-called plasmon pole approximation (PPA). In PPA, the inverse of dielectric function in Eq (2.100) is approximated with a single pole function[21]

$$[\varepsilon^{-1}(\mathbf{q},\omega)]^{\mathbf{G},\mathbf{G}'} \approx \delta_{\mathbf{G},\mathbf{G}'} + \left[\frac{R^{\mathbf{G},\mathbf{G}'}(\mathbf{q})}{\omega-\Omega_{\mathbf{G}\mathbf{G}'}(\mathbf{q})+i\delta}-\frac{R^{\mathbf{G},\mathbf{G}'}(\mathbf{q})}{\omega+\Omega_{\mathbf{G}\mathbf{G}'}(\mathbf{q})-i\delta}\right], \tag{2.103}$$

and the residuals $R^{\mathbf{G},\mathbf{G}'}$ and energies $\Omega_{\mathbf{G}\mathbf{G}'}$ are found by imposing the PPA to reproduce the exact ε^{-1} function.

Once that $\Sigma_{n\mathbf{k}}^{\mathrm{GW}}(\omega)$ is calculated, the quasiparticle energies $\varepsilon_{n\mathbf{k}}^{\mathrm{GW}}$ are obtained in QPA according to Eq (2.78).

$$\varepsilon_{n\mathbf{k}}^{\mathrm{GW}} = \varepsilon_{n\mathbf{k}}^{\mathrm{KS}} + Z_{n\mathbf{k}}\left[\Sigma_{n\mathbf{k}}^{\mathrm{GW}}\left(\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}\right) - V_{n\mathbf{k}}^{\mathrm{XC}}\right], \tag{2.104}$$

where $V_{n\mathbf{k}}^{\mathrm{XC}}$ is the exchange-correlation potential for KS orbital (n,$\mathbf{k}$). Its subtraction is required in order to avoid the double count of the exchange correlation interactions.

Using $\varepsilon_{n\mathbf{k}}^{\mathrm{GW}}$ in Eq (2.104), one can update $G_0(\mathbf{r},\mathbf{r}',\omega)$ in Eq (2.93) and $P_0(\mathbf{r},\mathbf{r}',\omega)$ in Eq (2.94) and calculate the self-energy and the quasiparticle energy iteratively until $\varepsilon_{n\mathbf{k}}^{\mathrm{GW}}$ is converged. Such *self-consistent* GW calculation, however, is known to overestimate bandwidths and bandgaps.[22] Practically, $\varepsilon_{n\mathbf{k}}^{\mathrm{GW}}$ with one single update of Eq (2.104), i.e., "one-shot" G_0W_0, or those obtained via self-consistent procedure of only the Green's function but not the screened interaction, i.e., GW_0, show better agreement with experiment.

As summarized in Ref [23], the best degree of self-consistency depends on the choice of exchange correlation functional in KS-DFT. For example, bandgaps at LDA+G_0W_0 level are known to be still underestimated while PBE0+G_0W_0 ones result overestimated. Both GGA+GW_0 and HSE03+G_0W_0 provide bandgaps close to experiment with similar accuracy.

2-4. Bethe-Salpeter Equation

GW approximation provides accurate quasiparticle gaps comparable with photoemission experiments. In photoabsorption experiment we have to consider e-h interaction since the number of electrons is fixed in the system. BSE is one of the approaches to treat such excited-state property. In this section theory of optical absorption and derivation of BSE from Hedin's equations are described.

2-4-1. Optical absorption spectrum

In solids, the optical absorption spectrum is defined as the imaginary part of the *macroscopic* dielectric function $\mathrm{Im}[\varepsilon_{\mathrm{M}}(\omega)]$. Using microscopic dielectric function defined in Eq (2.98), $\varepsilon_{\mathrm{M}}(\omega)$ is expressed in the long wavelength limit as

$$\varepsilon_{\mathrm{M}}(\omega) = \lim_{\mathbf{q}\to 0} \frac{1}{[\varepsilon^{-1}(\mathbf{q},\omega)]^{\mathbf{G}=0,\mathbf{G}'=0}(\mathbf{q},\omega)}, \tag{2.105}$$

Note that $\varepsilon_{\mathrm{M}}(\omega)$ is not same as $\lim_{\mathbf{q}\to 0} \varepsilon^{\mathbf{G}=0,\mathbf{G}'=0}(\mathbf{q},\omega)$ unless ε is a diagonal matrix. The difference $\varepsilon_{\mathrm{M}}(\omega) - \lim_{\mathbf{q}\to 0} \varepsilon^{\mathbf{G}=0,\mathbf{G}'=0}(\mathbf{q},\omega)$ is known as *local field effect*, which is not negligible in the inhomogeneous system. By inversion of $\varepsilon(\mathbf{q},\omega)$, Eq (2.105) can be rewritten as

$$\varepsilon_{\mathrm{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0} v_0(\mathbf{q}) \iint d\mathbf{r} d\mathbf{r}' \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{r}-\mathbf{r}')} \bar{P}(\mathbf{r},\mathbf{r}';\omega) \tag{2.106}$$

with $v_{\mathbf{G}}(\mathbf{q}) = \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2}$ being the Fourier transformed Coulomb potential and $\bar{P}(\mathbf{r},\mathbf{r}';\omega)$ the modified polarizability. $\bar{P}$ is obtained through Dyson-like equation[14],

$$\bar{P} = P + P\bar{v}\bar{P}, \tag{2.107}$$

where $\bar{v}$ is the modified Coulomb interaction $\bar{v}_{\mathbf{G}}(\mathbf{q})$ expressed as

$$\bar{v}_{\mathbf{G}}(\mathbf{q}) = \begin{cases} 0 & \mathbf{G} = 0 \\ \dfrac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} & \mathbf{G} \neq 0 \end{cases}. \tag{2.108}$$

From Eq (2.106) and (2.107) it is straightforward that the optical absorption can be calculated once polarizability P is obtained.

2-4-2. Polarizability by Bethe-Salpeter equation

We already know that RPA is the approximation of $P = P_0$. Optical absorption spectrum at RPA level stems from the collection of independent-particle excitations, thus RPA is also called independent-particle (IP) approximation. In this approximation, only exchange e-h interaction is included through local field effect. Furthermore, attractive e-h interaction is introduced by solving BSE.

As shown in subsection 2-3-1, the simplest approximation on Hedin's equations from Eq (2.85) to (2.89) is GW approximation. BSE is obtained as the second iteration of Hedin's equations. Substituting Eq (2.89) into Eq (2.86), we obtained an updated vertex function,

$$\Gamma(123) = \delta(12)\delta(13) + iW(12)\int G(16)G(72)\Gamma(673)\, d(67), \tag{2.109}$$

where we have assumed the approximation $\frac{\delta\Sigma(12)}{\delta G(45)} \approx iW(12)\delta(45)$.

Once we define three-point polarizability as follows

$$^3P(123) = -i\int G(16)G(72)\Gamma(673)\,d(67),$$
$$^3P_0(123) = -iG(13)G(32), \tag{2.110}$$

by multiplying $-iG(41)G(25)$ on the left side of Eq (2.109) and integrating over $d(12)$, we get an integral equation of polarizability,

$$^3P(453) = {}^3P_0(453) + i\int G(41)G(25)W(12)\,{}^3P(123)\,d(12). \tag{2.111}$$

Furthermore, a general four-point polarization is introduced as follows,

$$^4P_0(1234) = -iG(13)G(42), \tag{2.112}$$

thus three-point polarization is rewritten as $^3P_0(123) = -iG(13)G(32) = {}^4P_0(1233)$. Using the four-point form, finally Eq (2.111) can be extended to

$$^4P(1234) = {}^4P_0(1234) - \int {}^4P_0(1256)W(56)\,{}^4P(5634)\,d(56). \tag{2.113}$$

Similarly, Eq (2.107) in two-point form can be also extended to four-point form;

$$^4\bar{P}(1234) = {}^4P(1234) + \int {}^4P(1255)\bar{v}(56)\,{}^4\bar{P}(6634)\,d(56). \tag{2.114}$$

Combining Eq (2.113) and (2.114), finally we obtain

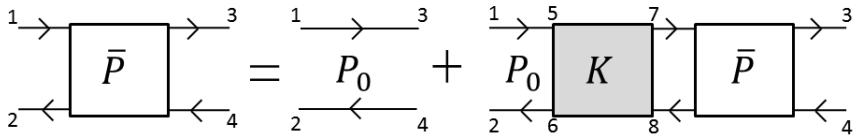
$$^4\bar{P}(1234) = {}^4P_0(1234) + \int {}^4P_0(1256)K(5678)\,{}^4P(7834)\,d(5678), \tag{2.115}$$

where the kernel *K* contains both exchange and attraction interactions as

$$K(5678) = \bar{v}(57)\delta(56)\delta(78) - W(56)\delta(57)\delta(68). \tag{2.116}$$

Eq (2.115) is BSE for the polarizability. In Fig. 2-6, BSE is represented using Feynman diagram.

$$\bar{P}\,(1234) = P_0(1234) + \int P_0(1256)K(5678)\bar{P}(7834)d(5678)$$



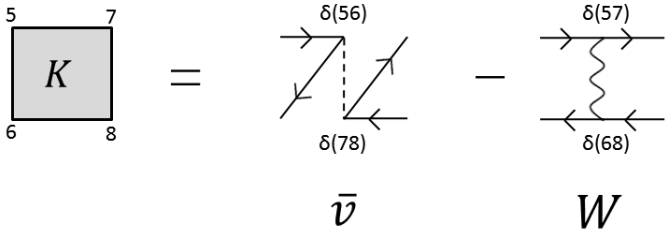


Figure 2-6. Feynman diagram of BSE

2-4-3. Practical calculation

The BSE for the polarizability can be expanded by introducing a single-particle Bloch basis set ($\{\psi_{n\mathbf{k}}\}$);

$$\begin{aligned}\bar{P}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4;\omega) = 2\sum_{n_1\mathbf{k}_1}\sum_{n_2\mathbf{k}_2}\sum_{n_3\mathbf{k}_3}\sum_{n_4\mathbf{k}_4}\psi_{n_1\mathbf{k}_1}(\mathbf{r}_1)\psi^*_{n_3\mathbf{k}_3}(\mathbf{r}_3)\psi_{n_4\mathbf{k}_4}(\mathbf{r}_4)\psi^*_{n_2\mathbf{k}_2}(\mathbf{r}_2)\\ \times \bar{P}_{(n_1\mathbf{k}_1 n_2\mathbf{k}_2)(n_3\mathbf{k}_3 n_4\mathbf{k}_4)}(\omega)\end{aligned} \tag{2.117}$$

We have interest in the case of $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ and $\mathbf{r}_3 = \mathbf{r}_4 = \mathbf{r}'$. In analogy with Eq (2.94) $\bar{P}_{(n_1\mathbf{k}_1 n_2\mathbf{k}_2)(n_3\mathbf{k}_3 n_4\mathbf{k}_4)}(\omega)$ is physically meaningful only when both $(n_1 n_2)$ and $(n_3 n_4)$ are pairs containing one conduction and one valence Bloch states.[14] In Tamm-Dancoff approximation, the coupling parts $(n_1 n_2 n_3 n_4) = \{vcc'v'\}$ and $(n_1 n_2 n_3 n_4) = \{cvv'c'\}$ are neglected. Finally, we only consider the resonant part $(n_1 n_2 n_3 n_4) = \{vcv'c'\}$.

In this basis, $\bar{P}$ is rewritten as

$$\bar{P}(\mathbf{r},\mathbf{r}';\omega) = -2i \sum_{v_1\mathbf{k}_1}\sum_{c_1\mathbf{k}_2}\sum_{v_2\mathbf{k}_3}\sum_{c_2\mathbf{k}_4} \psi_{v_1\mathbf{k}_1}(\mathbf{r})\psi^*_{v_2\mathbf{k}_3}(\mathbf{r}')\psi_{c_2\mathbf{k}_4}(\mathbf{r}')\psi^*_{c_1\mathbf{k}_2}(\mathbf{r}) \times L_{(v_1\mathbf{k}_1 c_1\mathbf{k}_2)(v_2\mathbf{k}_3 c_2\mathbf{k}_4)}(\omega), \tag{2.118}$$

where $L = i\bar{P}$ is electron-hole (two-particle) Green's function. From Eq (2.106) and (2.118), optical absorption spectrum is represented as

$$\mathrm{Im}[\varepsilon_{\mathrm{M}}(\omega)] = 2\lim_{\mathbf{q}\to 0} v(\mathbf{q}) \times \sum_{\mathbf{K}_1}\sum_{\mathbf{K}_2} \langle\psi_{v_1\mathbf{k}_1}|e^{-i\mathbf{q}\mathbf{r}}|\psi_{c_1\mathbf{k}_1+\mathbf{q}}\rangle\langle\psi_{c_2\mathbf{k}_2+\mathbf{q}}|e^{i\mathbf{q}\mathbf{r}'}|\psi_{v_2\mathbf{k}_2}\rangle \mathrm{Re}\big[L_{\mathbf{K}_1\mathbf{K}_2}(\omega)\big], \tag{2.119}$$

thus now the remaining observable to calculate is the interacting electron-hole Green's function $L_{\mathbf{K}_1\mathbf{K}_2}(\omega)$. Here **K** represents the general conduction-valence pairs, $\mathbf{K}_i = (c_i, v_i, \mathbf{k}_i)$. From BSE for the polarizability (Eq (2.115)), we can redefine BSE for $L_{\mathbf{K}_1\mathbf{K}_2}(\omega)$:

$$L_{\mathbf{K}_1\mathbf{K}_2}(\omega) = L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega)\left(\delta_{\mathbf{K}_1\mathbf{K}_2} + \sum_{\mathbf{K}_3}\Xi_{\mathbf{K}_1\mathbf{K}_3}(\omega)L_{\mathbf{K}_3\mathbf{K}_2}(\omega)\right), \tag{2.120}$$

where the Bethe-Salpeter kernel $\Xi = -iK$ is defined as $\Xi_{\mathbf{K}_1\mathbf{K}_2} = -i\bar{v}_{\mathbf{K}_1\mathbf{K}_2} + iW_{\mathbf{K}_1\mathbf{K}_2}$. BSE results similar to Dyson's equation for one-particle Green's function (Eq (2.73)).

$L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega)$ is the free electron-hole Green's function represented as

$$L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega) = i\left[\frac{1}{\omega - \varepsilon_{c_1\mathbf{k}_1} + \varepsilon_{v_1\mathbf{k}_1} + i\delta} - \frac{1}{\omega + \varepsilon_{c_1\mathbf{k}_1} - \varepsilon_{v_1\mathbf{k}_1} - i\delta}\right]\delta_{\mathbf{K}_1\mathbf{K}_2}. \tag{2.121}$$

The exchange interaction $\bar{v}_{\mathbf{K}_1\mathbf{K}_2}$ is defined as

$$\bar{v}_{\mathbf{K}_1\mathbf{K}_2} = \frac{1}{\Omega}\sum_{\mathbf{G}\neq 0}\langle\psi_{c_2\mathbf{k}_2}|e^{i\mathbf{G}\mathbf{r}}|\psi_{v_2\mathbf{k}_2}\rangle\langle\psi_{v_1\mathbf{k}_1}|e^{-i\mathbf{G}\mathbf{r}'}|\psi_{c_1\mathbf{k}_1}\rangle v^{\mathbf{G}}. \tag{2.122}$$

Similarly, direct electron-electron scattering term $W_{\mathbf{K}_1\mathbf{K}_2}$ is written as follows

$$W_{\mathbf{K}_1\mathbf{K}_2}(\omega) = \frac{1}{\Omega}\sum_{\mathbf{G},\mathbf{G}'}\langle\psi_{c_2\mathbf{k}_2}|e^{i(\mathbf{k}_2-\mathbf{k}_1+\mathbf{G})\mathbf{r}}|\psi_{c_1\mathbf{k}_1}\rangle\langle\psi_{v_1\mathbf{k}_1}|e^{-i(\mathbf{k}_2-\mathbf{k}_1+\mathbf{G}')\mathbf{r}'}|\psi_{v_2\mathbf{k}_2}\rangle W^{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega). \quad (2.123)$$

The dynamical screened interaction W is often approximated to be static $\omega = 0$. It is justified since the plasma frequency of the investigated system is usually larger than the excitonic binding energies.[14]

Using Eq (2.121), (2.122), and (2.123), BSE can be solved. After some algebraic transformations, BSE in Eq (2.120) can be also represented as

$$-iL_{\mathbf{K}_1\mathbf{K}_2}(\omega) = [H^{exc} - I\omega]^{-1}_{\mathbf{K}_1\mathbf{K}_2}, \quad (2.124)$$

where $I = \delta_{\mathbf{K}_1\mathbf{K}_2}$ and two-particle excitonic Hamiltonian, $H^{exc}_{\mathbf{K}_1\mathbf{K}_2}$ is defined as

$$H^{exc}_{\mathbf{K}_1\mathbf{K}_2} = (\varepsilon_{c_1\mathbf{k}_1} - \varepsilon_{v_1\mathbf{k}_1})\delta_{\mathbf{K}_1\mathbf{K}_2} - i\Xi_{\mathbf{K}_1\mathbf{K}_2}. \quad (2.125)$$

Now, diagonalization of $H^{exc}_{\mathbf{K}_1\mathbf{K}_2}$ provides excitonic states stemming from mixing of independent-particle excitations. Using the eigenvalues E_λ and eigenvectrors $|\lambda\rangle = \sum_i A^{\mathbf{K}_i}_\lambda \psi_{v_i\mathbf{k}_i}(\mathbf{r})\psi^*_{c_i\mathbf{k}_i}(\mathbf{r}')$, Eq (2.124) is transformed into the spectral representation as

$$-iL_{\mathbf{K}_1\mathbf{K}_2}(\omega) = \sum_\lambda \frac{A^{\mathbf{K}_1}_\lambda A^{*\mathbf{K}_2}_\lambda}{E_\lambda - \omega + i\delta}. \quad (2.126)$$

Finally, the optical absorption spectrum in Eq (2.119) is represented as follows.

$$\mathrm{Im}[\varepsilon_{\mathrm{M}}(\omega)] = 2\lim_{\mathbf{q}\to 0} v(\mathbf{q}) \times \sum_{\mathbf{K}_1}\sum_{\mathbf{K}_2}\sum_\lambda \langle\psi_{v_1\mathbf{k}_1}|e^{-i\mathbf{q}\mathbf{r}}|\psi_{c_1\mathbf{k}_1+\mathbf{q}}\rangle\langle\psi_{c_2\mathbf{k}_2+\mathbf{q}}|e^{i\mathbf{q}\mathbf{r}'}|\psi_{v_2\mathbf{k}_2}\rangle A^{\mathbf{K}_1}_\lambda A^{*\mathbf{K}_2}_\lambda \delta(\omega - E_\lambda). \quad (2.127)$$

Diagonalization of $H^{exc}_{\mathbf{K}_1\mathbf{K}_2}$ is computationally demanding with respect to the storage in memory and the calculation time due to the large size of excitonic Hamiltonian ($N_v \times N_c \times N_{\mathbf{k}}$). One of the approaches to solve the problem is Lanczos-Haydock recursion method. For the detailed description, see Ref [21].

Bibliography

(1) Martin, R. M. *Electronic structure: Basic theory and practical methods*; Cambridge University Press: Cambridge, UK, 2004.

(2) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(3) Levy, M. *Phys. Rev. A* **1982**, *26*, 1200.

(4) Lieb, E. H. *Int. J. Quantum Chem.* **1983**, *24*, 243.

(5) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(6) Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980**, *45*, 566.

(7) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048.

(8) Perdew, J. *Phys. Rev. Lett.* **1985**, *55*, 1665.

(9) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(10) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(11) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

(12) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207.

(13) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2006**, *124*, 219906.

(14) Onida, G.; Reining, L.; Rubio, A. *Rev. Mod. Phys.* **2002**, *74*, 601.

(15) Mahan, G. D. *Many-Particle Physics*; 2nd ed.; Plenum Press: New York, 1990.

(16) Haug, H. J. W.; Jauho, A.-P. *Quantum Kinetics in Transport and Optics of Semiconductors*; 2nd ed.; Springer: Berlin; New York, 2008.

(17) Mattuck, R. D. *A Guide to Feynman Diagrams in the Many-Body Problem*; 2nd ed.; McGraw-Hill: New York, 1976.

(18) Fetter, A. L.; Walecka, J. D. *Quantum Theory of Many-Particle Systems*; Dover Publications: New York, 2003.

(19) Aryasetiawan, F.; Gunnarsson, O. *Reports Prog. Phys.* **1998**, *61*, 237.

(20) Shishkin, M.; Kresse, G. *Phys. Rev. B* **2006**, *74*, 035101.

(21) Marini, A.; Hogan, C.; Grüning, M.; Varsano, D. *Comput. Phys. Commun.* **2009**, *180*, 1392.

(22) Shishkin, M.; Kresse, G. *Phys. Rev. B* **2007**, *75*, 235102.

(23) Fuchs, F.; Furthmüller, J.; Bechstedt, F.; Shishkin, M.; Kresse, G. *Phys. Rev. B* **2007**, *76*, 115109.

Chapter 3.
Bandgap and Band Bending of Water-splitting Photocatalyst GaN:ZnO

3-1. Introduction to GaN:ZnO

3-1-1. Experimental facts

A more performing, clean, and sustainable sun-to-energy production of hydrogen requires improved efficiencies of water-splitting photocatalyst. After the discovery of Honda-Fujishima effect,[1] many metal oxides have been suggested as promising water-splitting photocatalysts.[2] However, the absorption edges of most of them lie in the ultraviolet-light wavelengths, thus only a few part of solar spectrum is converted into chemical energy. Accordingly, (oxy)nitrides have been widely investigated in recent years because the higher potential energy of N 2p orbital than O 2p ones leads to higher valence top, which results in narrower bandgaps in visible light absorption.[3,4]

Maeda and Domen reported gallium zinc oxynitride (GaN:ZnO) as the first successful example of overall water-splitting photocatalyst.[5–7] GaN:ZnO with the Zn concentration less than 22 % absorbs visible light at wavelengths in the range 400-500 nm. Moreover, the band

edge position meets thermodynamic requirement for overall water splitting reaction: the potential of conduction band minimum (CBM) is more negative than the redox potential of H^+/H_2 and that of the valence band maximum (VBM) is more positive than that of O_2/H_2O.[8] Satisfying this ideal electronic requirement, GaN-rich solid solutions enable water-splitting reaction under visible-light irradiation. The quantum efficiency has been reached to about 5.9 % at 420-440 nm modified with Rh-Cr mixed oxide cocatalyst.[9] More recently, long-time operation test showed that the solid solution remains active for 3 months without noticeable degradation.[10] Owing to the combination of visible-light response, high activity, and high durability, GaN:ZnO is undoubtedly one of the most promising water splitting photocatalysts for massive hydrogen production.

GaN:ZnO has wurtzite structure ($P6_3mc$) same as pure GaN and ZnO parental compounds. Energy dispersive X-ray spectroscopy (EDX) indicates the atomic ratio of Ga to N and Zn to O is close to 1.[7] Thus, this solid solution is regarded as a pseudo binary III-V/II-VI alloy, i.e., $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ with the concentration of ZnO x. There is no report that pure compounds GaN and ZnO are good water-splitting photocatalysts. Although the band edge positions of GaN and ZnO also satisfy the thermodynamic requirement,[8,11,12] their bandgaps are too wide to absorb visible light, i.e., 3.4 eV and 3.2 eV, respectively. Under UV irradiation, pure GaN with high crystallinity is reported to have photocatalytic activity, but the

quantum efficiency is lower than that of GaN:ZnO.[13] These facts indicate GaN:ZnO has several advantages due to the chemical alloying between GaN and ZnO.

In the works by Maeda and Domen, the solid solutions have been intensively studied in GaN-rich regime because their synthetic method, the nitridation of the mixture of β-Ga_2O_3 and ZnO under NH_3 flow, is limited to GaN-rich case with $x < 0.42$.[6,7] First report of synthesis of ZnO-rich alloys is the solid-state reaction between ZnO and GaN powder under high pressure and high temperature (HPHT).[14] This method enables to synthesize wide range of compositions from GaN-rich to ZnO-rich solid solutions. The ZnO-rich sample from this method showed an absorption edge longer than 500 nm i.e., the bandgap is less than 2.5 eV. The bandgap is clearly narrower than those of GaN-rich ones (~ 500 nm) in the range x = 0.03-0.13.[5] Similarly, ZnO-rich samples obtained via the solution combustion method,[15] the nitridation of layered double hydroxides (LDHs),[16] and the nitridation of the mixture of $ZnO/ZnGa_2O_4$ nanocrystals[17] also showed band gap less than 2.5 eV for $x > 0.8$.

From these reports, ZnO-rich GaN:ZnO is expected to be more promising than GaN-rich ones with respect to the visible-light response. Nevertheless, there is still no report to show overall water-splitting reaction on ZnO-rich samples. Anyway, the origin of such larger bandgap lowering in ZnO-rich alloys has not been revealed yet and hence further investigation is still required.

3-1-2. Issues to investigate chemical alloying effect

The mechanism of bandgap narrowing by chemical alloying is still a controversial issue. Previous theoretical and experimental works suggested the mechanism of bandgap lowering is N 2p-Zn 3d repulsion that shifts the VBM upward.[18–21] On the other hand, photoluminescence spectroscopy and first-principles calculation by Yoshida *et al.*[22] proposed an impurity-like picture where the visible-light absorption is attributed to the transition from Zn acceptor levels to CBM of GaN. In contrast, soft X-ray spectroscopic studies attributed the bandgap reduction in the solid solution with x = 0.42 to the repulsion of CB edges of GaN phase and ZnO phase.[23]

Although much literature focuses on $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ from the theoretical point of view by means of first-principles calculations, the mechanism of bandgap lowering has not been clarified due to many complicated factors. The main three issues involved are presented in the following.

1. Atomic configuration

The first and the most difficult to tackle is the modellization of atomic configuration in the crystal, i.e., randomness or short/long-range ordering. The XRD analyses in all of the experimental reports indicate these solid solutions form single wurtzite phase. However, the distribution of ions in the parental lattice sites is still unclear and it is likely to depend on the synthetic conditions. Such *atomic configuration* has striking effect because it determines the

character of local chemical bonding, which is important in materials with covalent character. $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ is one of *non-isovalent* alloys,[24,25] where the octet rule violation exists inevitably due to the mismatch of valence electrons, i.e., GaN is III-V compound and ZnO is II-VI compound. Such non-isovalent character requires a careful treatment of the alloying structure modellization, even for fixed ZnO concentration regime.

In the first theoretical works of GaN:ZnO, the alloying structures have been modelized by means of special quasirandom structure (SQS) approach in the assumption that the alloy is a random solution.[18] Afterward, however, Monte Carlo (MC) simulations showed that there is strong short-range order (SRO), which leads to clustering of the dilute compound in the matrix of the rich compound.[26,27] DFT calculations on the MC modelized structures at x = 0.2, 0.5, and 0.8 at T = 1200K predict symmetric bandgap, which is incompatible with the dramatic bandgap reduction shown in experiments for ZnO-rich samples.[14–17] It indicates thermodynamically stable configuration is not experimentally achieved in the ZnO-rich solid solutions. Therefore, further detailed investigation on the atomic configurations is required to describe the mechanism of bandgap reduction in GaN:ZnO.

2. Band edge position

The second issue is that exact energies of valence and conduction band edges vs vacuum level, i.e., band edge positions, are not obtained from standard bulk calculations employing periodic

boundary conditions. This is due to the contribution of zero momentum components of pseudopotentials.[28,29] As a result, calculations on bulk systems can describe bandgap change by chemical alloying, but not the band edge shift. It hinders us to ascribe the origin of bandgap narrowing to VBM raising or CBM lowering.

One of the methods to solve this problem is to calculate electrostatic potentials on the slab models, which has 2-dimensional periodic surface with thick vacuum region, or heterostructure having interface with a different compound.[30,31] Some literatures reported on calculations of slab models of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$,[32,33] however, their discussions are still not satisfactory to reveal the mechanism of bandgap narrowing. This is because the impact of the atomic configuration is not considered carefully, thus there is still need for further investigation.

3. Accuracy of bandgap calculation

The third is the accuracy of DFT calculations in predicting bandgaps. Although LDA and GGA are the common techniques for Kohn-Sham DFT calculations for condensed matter, these approximations fail to describe strongly correlated electronic systems properly. Thus DFT calculations show errors on $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$, in particular on ZnO-rich alloys since shallow Zn 3d level plays a crucial role in the valence band.[34] On the other hand, calculation on GaN-rich one is not so problematic since the impact of deep Ga 3d level is not significant

on the character of frontier orbitals. To overcome this problem, DFT+U approach,[35] which adds onsite Coulomb interactions on *d* electrons,[18] and hybrid exchange correlation (XC) functional as B3LYP or HSE[21,33] have been employed. Nevertheless, the accuracy of bandgap is still biased due to the fact that DFT is a theory for ground state. Empirical LDA+C[26] and composition-dependent adjustments[27] were also employed but the quantitative reliability have to be considered carefully. Therefore, more rigorous *ab initio* approach is desired for the quantitative comparison with the experimental results.

MBPT scheme based on GW approximation for quasi-particle gaps and Bethe-Sapleter equation (BSE) for the electron-hole interactions have been revealed to be a powerful tool over the last decades.[36] Previous literature reported a combined GW+BSE calculation on $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ systems, but their discussion on the mechanism of the visible-light response was still not satisfactory for discussing the mechanism of bandgap narrowing.[37]

3-1-3. Objective of this work

In the following sections, I tackle the three issues described in subsection 3-1-2. In Sec 3-2, at first, different levels of calculations, LDA, LDA+U, and GW approximation, are checked for pure GaN and ZnO in terms of the accuracies for ground-state electronic structures and bandgaps. I will show that GW approximation perfectly reproduces the bandgaps of both GaN and ZnO. For GaN, LDA+U level of theory is also a feasible and reliable approach.

In Sec 3-3, GaN-rich slab models are investigated by means of LDA+U to reveal the band edge position shifts of VBM and CBM caused by ZnO doping. The impact of the atomic configurations on the band edge positions is discussed. As secondary product, I found the band bending at the surface stemming from the electric field by the opposite charges on donor O and accepter Zn.

In Sec 3-4, my focus moves on the accuracy and comparison with experiment. GW approximation is employed to obtain accurate bandgaps. Moreover, absorbance is also computed by solving BSE, which can be directly compared with the experimental UV-Vis diffuse reflectance spectra (DRS). From the agreement with experiment I suggest that metastable structure with large number of Zn-N bonds is expected in ZnO-rich alloys.

3-2. Preliminary calculations for pure GaN and ZnO

As preliminary calculations, bandgaps of both pure GaN and ZnO bulks are calculated at LDA+U, PBE+G_0W_0, and HSE06+G_0W_0 levels and their results are compared to check the accuracy of calculations. At first, the geometries are optimized by LDA calculations. Grids of Γ-centered k-points sampling were set to 10 × 10 × 10. Structural optimization was terminated when the forces on all of the atoms were smaller than 10 meV/Å. The resulting lattice parameters of GaN (ZnO) are a = 3.14 (3.17) Å and c = 5.11 (5.12) Å, which are in agreement with the experimental values a = 3.19 (3.25) Å and c = 5.19 (5.21) Å.[38,39] Then, on the relaxed structures, LDA+U,[35] PBE+G_0W_0,[40] and HSE06+G_0W_0[41] calculations were performed to obtain an improved description of the electronic properties. Tab. 3-1 shows the resulting bandgaps E_g and the positions of cation 3d bands from the top of valence bands.

Table 3-1. Bandgaps and levels of d bands of GaN and ZnO

		LDA	LDA+U	PBE+G_0W_0	HSE06+G_0W_0	Exp
GaN	E_g /eV	2.23	3.02	3.48	4.05	3.507[38]
	d band /eV	13.2	17.1	16.2	16.9	17.1[42]
ZnO	E_g /eV	0.84	1.84	2.55	3.41	3.44[43]
	d band /eV	4.6	7.1	6.0	6.5	6.95[44], 7.4[45]

As well known, LDA calculations underestimate bandgaps due to the self-interaction error[46,47] and derivative discontinuity of the XC energy.[48,49] To remove the self-interaction

error, LDA+U calculations with on-site Coulomb interactions U_{eff} of 8.5 eV and 8.0 eV for 3d orbitals of Ga and Zn respectively are applied. These values provide good agreement between theoretically calculated and experimentally reported 3d bands of Ga (Zn) in GaN[42] (ZnO[44,45]). Thus the resulting electronic structures of the ground states are largely improved. Despite an improvement in the description of d bands, band gaps E_g by LDA+U are still underestimated in particular for ZnO. As mentioned above, such severer error comes from the fact that shallow Zn 3d level plays a crucial role in the valence band.[34]

Strictly speaking, bandgaps measured by experimental techniques involving electronic excitations are quasiparticle gaps, therefore Kohn-Sham (KS) gaps from DFT calculations for ground states are not comparable. Quasiparticle gaps can be addressed by GW approximation which takes dynamical screening potential into account. I employ "one-shot" GW scheme (G_0W_0),[40,41] being the resulting bandgaps governed from the initial KS orbitals. Two XC functionals, the GGA functional of Perdew-Burke-Ernzerhof (PBE)[50] and the hybrid functional by Heyd-Scuseria-Ernzerhof (HSE),[51,52] are employed.

The calculation at the PBE+G_0W_0 level of theory reproduces the experimentally measured bandgap of GaN well, while it still underestimates that of ZnO. On the other hand, HSE06+G_0W_0 calculations show good agreement with the experimental data for the bandgap of ZnO, but open the gap of GaN too much. Consistently with my calculations, an

overestimated value was also reported for GaN by HSE06+G_0W_0 in previous literature[53]. My HSE06+G_0W_0 calculations give excellent agreement with the experiment for ZnO, but is slightly larger than the previous reported value 3.21 eV by HSE03+G_0W_0.[54] This is ascribed to the fact that the bandgap in the literature is calculated on top of the GGA optimized geometry, where overestimation of the lattice parameter occurs, resulting in a smaller bandgap. The results in Table 3-1 indicate different approaches should be applied for GaN:ZnO solid solutions depending on the ratio of GaN to ZnO: PBE+G_0W_0 suitable for GaN-rich alloys, while HSE06+G_0W_0 is suitable for ZnO-rich ones.

Although GW approximation provides accurate bandgaps, the computational burden is quite larger than ones of DFT calculations, thus its usage in large systems with more than one-hundred atoms results undoable. On the other hand, the computational cost of LDA+U is almost same as normal LDA calculation, thus it is easily applied to large systems. As shown in Tab. 3-1, LDA+U gives meaningful results with reduced efforts in the case of GaN. From this preliminary analysis, in the following I opted for two approaches, i.e.: 1) slab models with more than one hundred atoms for GaN-rich solid solution (x = 0.0345) by LDA+U (Sec 3-3), and 2) smaller bulk models for both GaN-rich and ZnO-rich models (x = 0.125 and 0.875) by GW approximation and BSE (Sec 3-4).

3-3. Band-edge character of ZnO-doped GaN

3-3-1. Computational details

As mentioned in Sec 3-1, the exact band positions are not obtained from calculations on bulk models. Due to the contribution of zero momentum components of pseudopotentials, the reference of eigenvalues is not the vacuum level anymore.[28,29] Slab model calculation solves this problem since it provides an electrostatic potential of vacuum region, thus subtracting the potential energy from the band energies, band edge positions *vs* vacuum level are obtained.

In this section, I investigate the band-edge character and band bending of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$. As my slab systems have more than one hundred atoms, GW scheme results in the difficult applicability. Hence my choice is to study GaN-rich alloys by means of LDA+U approach. Here, only small ZnO concentrations less than 10% are considered. I employed Vienna Ab initio Simulation Package (VASP).[55] The projector-augmented wave (PAW) method with LDA was used with a cutoff energy of 550 eV.

I tested two nonpolar facets $(11\bar{2}0)$ and $(1\bar{1}00)$. In this work, polar (0001) and $(000\bar{1})$ facets were not considered because the complicated reconstruction of such surfaces is not suitable to focus on the effect of ZnO doping. My $(11\bar{2}0)$ and $(1\bar{1}00)$ models include the 29 and 16 (116 and 64 atoms) unit layers, respectively, to get comparable thicknesses for the two slabs. The thickness of the vacuum region is ≈10 Å. The lattice parameters are a = 5.43 Å, b =

5.11 Å, and c = 53.48 Å for the $(11\bar{2}0)$ model and a = 3.14 Å, b = 5.11 Å, and c = 51.46 Å for the $(1\bar{1}00)$ model. Grids of Γ-centered k-points sampling were set to 5 × 5 × 1 and 8 × 5 × 1, respectively. Structural optimization was terminated when the forces on all of the atoms were smaller than 50 meV/Å. All atoms were relaxed with fixed lattice constants.

The LDA+U calculated band plots are shown in Fig. 3-1. Red lines indicate the bands of surface states. Avoiding these bands, we can define the band-edge positions in the bulk region (blue dashed lines). My results show that the band-edge positions are identical in the two facets. The CBM and VBM of the $(11\bar{2}0)$ ($(1\bar{1}00)$) facet are at –6.54 (–6.55) eV and –3.46 (–3.48) eV from the vacuum level and the bandgaps have similar values to the bulk value (3.02 eV). This means both slab models are converged to reproduce the bulk's electronic structure and thus appropriate for use in investigating the band-edge position of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$.

It should be noted that the LDA+U-calculated bandgap still underestimates the experimental value of 3.507 eV.[38] Compared with previous first-principles calculations of band alignment,[56] the VBM is higher by about 0.17 eV and the CBM is lower by about 0.22 eV. This means that my calculation provides the correct center of the bandgap even though the bandgap is underestimated. It is consistent with the findings of a previous study,[28,48] thus offering a valuable starting point for my discussion.

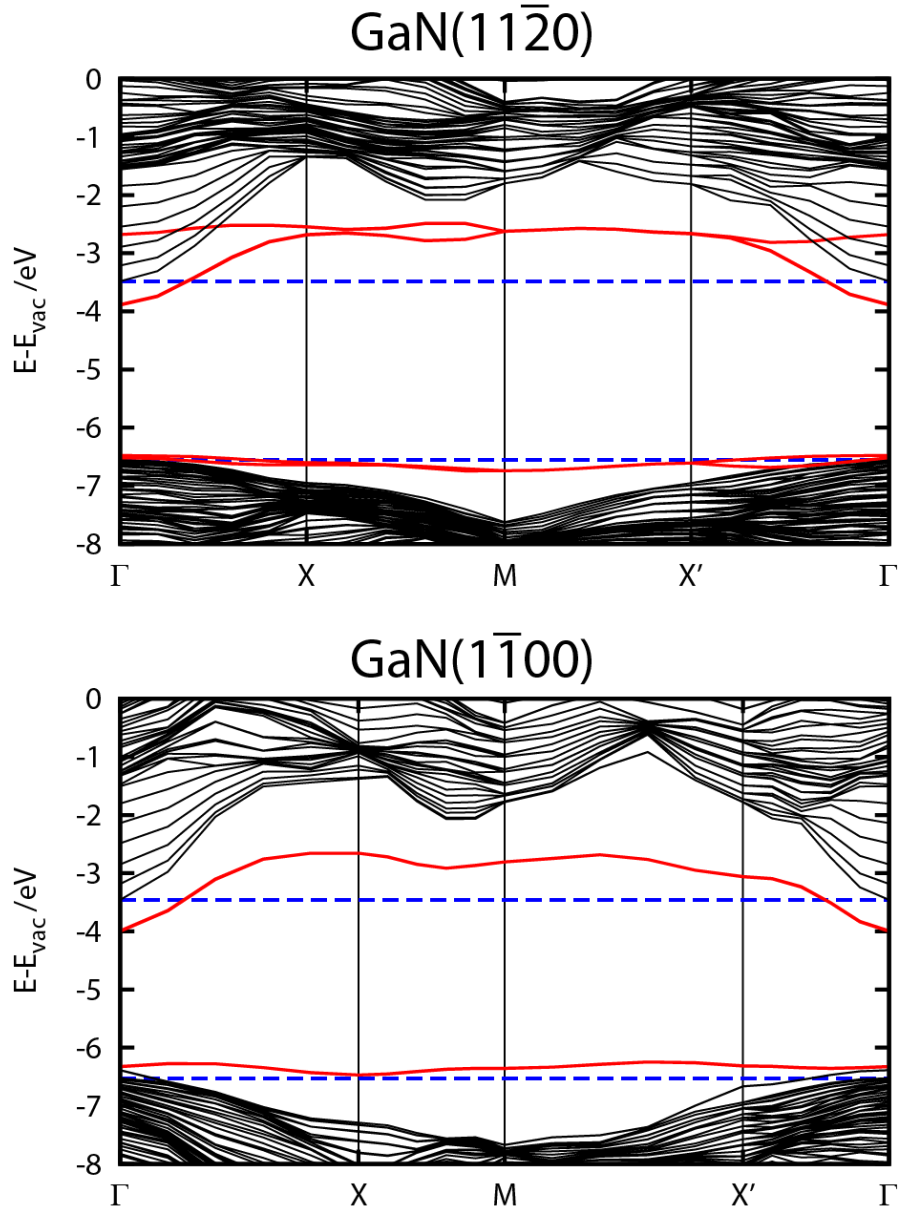


Figure 3-1. Band plots of (11$\bar{2}$0) (upper) and (1$\bar{1}$00) (lower) surfaces of GaN. Red lines are bands of surface states. Blue dashed lines show the band-edge positions of the bulk regions.

In this work, I analyze the ZnO-doping effect in GaN (11$\bar{2}$0), a facet that is characterized by the ABAB… stacking pattern. Its structure is shown in Fig. 3-2. I define the positions of

Zn and O dopant atoms by indexing the planes from 0 to 14, as shown in the top panel of Fig. 3-2. To avoid any net dipole vertical to the surface plane, Zn and O symmetrically dope both sides of the slab. Thus the indexes are symmetric. I only consider substitution of Ga and N with Zn and O, respectively. In the wurtzite structure, each cation and anion has four chemical bonds: three short equatorial bonds (1.918 Å, blue in Fig. 3-2) and one longer axial bond (1.924 Å, red in Fig. 3-2). Such anisotropic character affects the ZnO doping effect, as shown below.

Because $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ belongs to the class of non-isovalent III-V/II-VI alloys, non-isovalent bonds (chemical bond breaking the octet-rule with ±1 electron) like Ga-O (III-VI) or Zn-N (II-V) inevitably occur. There also could be other bonds with ±2 or ±3 electrons like Ga-Ga or O-N: I neglect them due to their larger thermodynamic instability. In addition, the experimentally reported nonstoichiometry, e.g. excess O element over N,[5] is also out of scope in this work.

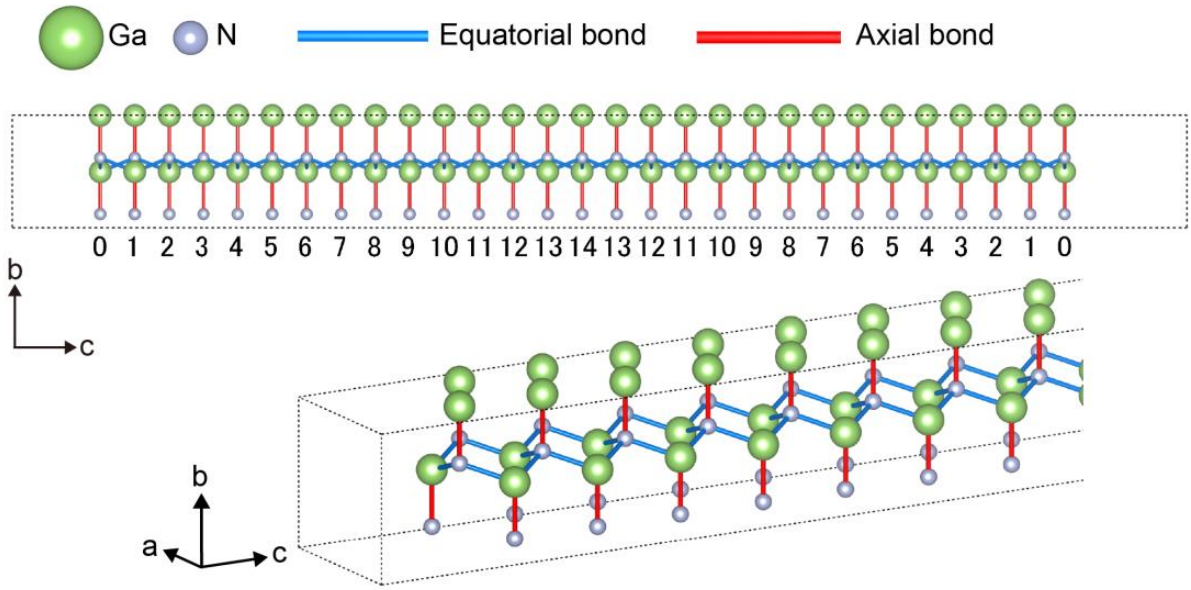


Figure 3-2. Slab model of the GaN $(11\bar{2}0)$ facet. Upper picture is viewed from the *a*-axis direction with the indexes of planes. The lower clearly shows that each N and Ga has one axial and three equatorial bonds.

3-3-2. Band-edge position

When discussing the band-edge character in the bulk region, effects of surface state should be in principle avoided. Therefore, the Zn and O atoms will dope deep layers far from the surface dangling bond. Moreover, doping different planes with Zn and O atoms induces a surface dipole as it will be shown in the next subsection; thus, in this subsection I only consider the case that the Zn and O atoms dope the same plane.

Fig. 3-3(a) shows the density of states projected on each plane (pDOSs) plotted along the plane-normal direction. These plots clearly show that the dangling-bond states lie at the surfaces (circled parts) and it is possible to define the band-edge positions of the bulk regions. Here, I compare GaN $(11\bar{2}0)$ and the ZnO-doped system where one Ga–N pair bonded axially in plane 9 is replaced by a Zn–O pair. This system is named “PairA-9”.

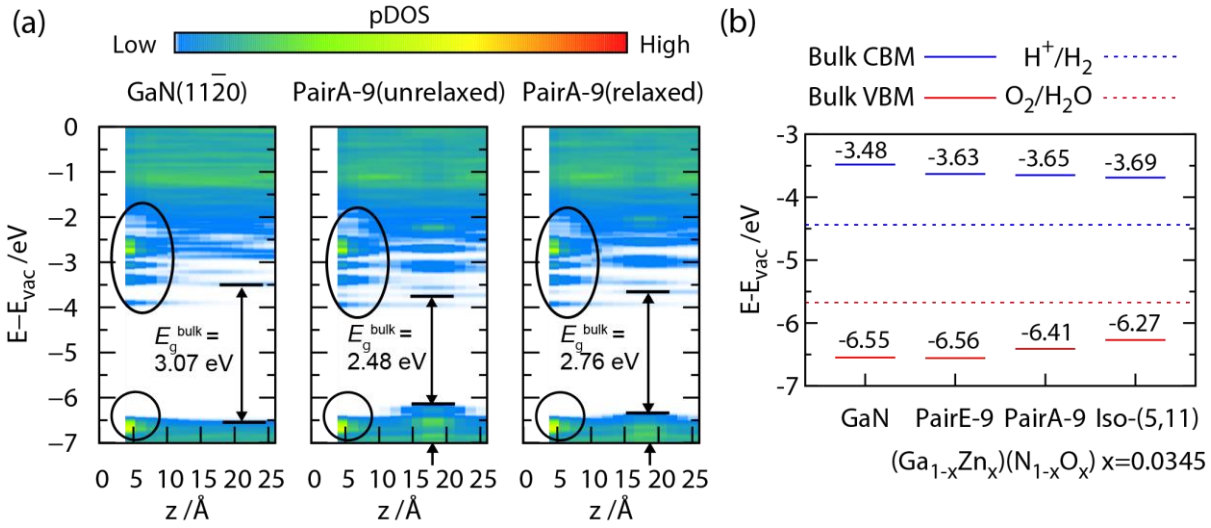


Figure 3-3. (a) pDOS plots along the plane-normal direction of GaN (11$\bar{2}$0) (left) and PairA-9 without and with structural relaxation (center and right). Surface states are circled and the bandgaps of the bulk regions are depicted. (b) Band-edge positions of the bulk region of GaN, PairE-9, PairA-9, and Iso-(5,11) models (CBM: blue, VBM: red). Dashed lines are redox potentials of H^+/H_2 (blue) and O_2/H_2O (red).

First, the unrelaxed PairA-9 model (Fig. 3-3(a) center) is compared with GaN (Fig. 3-3(a) left) to evaluate the bare substitution chemical effect. Obviously, the bandgap narrowing is caused by both CBM lowering and VBM raising. In particular the VBM raising is large. As reported in previous literature, such VBM raising is due to the anti-bonding orbital by the coupling of N 2p and Zn 3d orbitals, i.e. p–d repulsion.[7,34] Therefore, the pDOS of the valence band is clearly localized around plane 9 including substituted Zn (indicated by arrows). On

the other hand, the CBM is delocalized over the bulk region. This band is mainly constituted by N 3s, Ga 4s, and O 3s orbitals, while the CBM of GaN is mainly formed by N 3s and Ga 4s. The additional mixing of the O 3s orbital enhances the bonding character, resulting in the lower band-edge position compared with that of the pristine undoped GaN $(11\bar{2}0)$ surface.

However, such a noticeable VBM raising is less evident after structural relaxation (Fig. 3-3(a) right). This is clearly explained in terms of increase in the Zn–N bond length in the relaxed structure (from 1.925 Å to 1.956 Å). Such structural expansion is induced by the octet-rule violation of Zn–N and Ga–O bonds. Longer Zn–N bonds decrease the coupling between N 2p and Zn 3d, weakening the p–d repulsion and pushing the VBM down. On the other hand, such a trend is not observed on the CBM. The local structural expansion does not impact on the spatially delocalized CBM, as testified by its position which remains almost unaltered after the relaxation.

In Fig. 3-3(b), the band-edge positions of other ZnO-doped models are also compared. In this case, all of the considered structures are relaxed. The concentration of ZnO, x, is identical among these models (x = 0.0345), and hence the differences derive only from the atomic configuration of Zn and O.

As with the PairA-9 model, “PairE-9” has one Zn–O pair in plane 9, but the Zn–O pair forms an equatorial bond which is shorter than an axial one. Compared with PairA-9, the

CBM position is consistent while the VBM position is pushed down. As a result, the VBM positions of PairE-9 and GaN are almost the same, with the bandgap narrowing only ascribed to the CBM lowering. The VBM of PairE-9 lower than that of PairA-9 is similarly explained as I did from the comparison between unrelaxed and relaxed PairA-9 models. Namely, when Zn–O pair in the PairE-9 is equatorially bonded, the three remaining Zn–N bonds are one long axial and two short equatorial bonds. The axial Zn-N bond, which is 1.971 Å (the equatorial bonds are 1.956 Å), causes decreased p–d repulsion and lowering the VBM.

In the "Iso-(5,11)" model, the Zn and O do not form pairs and are isolated. To avoid Zn–O bonding, the unit cell was doubled along the *a*-direction and two planes (5 and 11) were doped to make the ZnO concentration x consistent with the other models. Iso-(5,11) is characterized by the most marked VBM raising in the three considered models. Again, this is explained by the Zn 3d –N 2p interaction. Zn atoms in Iso-(5,11) have four Zn–N bonds, while PairE-9 and PairA-9 have only three Zn–N bonds. The increased number of Zn–N bonds causes further p–d repulsion. This is the reason for the noticeable VBM raising. As expected, the CBM position does not differ from the other two ZnO-doped models because of its delocalized character again.

As mentioned, my calculations overestimate (underestimate) the VBM (CBM) position because of the well-known DFT shortcomings in predicting semiconductor bandgaps;

nevertheless, all models meet the requirement that the VBM (CBM) is lower (higher) than the redox potential of O_2/H_2O (H^+/H_2). This result shows that the $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ system is suitable for the overall water splitting, as is experimentally known.

Although the VBM raising of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ enhances visible-light absorption, the orbital is localized around Zn–N bonds as shown in Fig. 3-3(a). Thus photogenerated holes will be trapped around Zn–N bonds, possibly resulting in the hole mobility reduction with suppression of subsequent oxygen evolution reaction (OER). This point will be discussed in subsection 3-3-4.

In summary, it is found that 1) bandgap lowering comes from both VBM raising and CBM lowering and 2) the impact of local atomic configuration is mainly on the VBM in GaN-rich compounds (x = 0.0345).

3-3-3. Band bending at the surface

Next, models where the Zn and O atoms dope different planes were studied. A previous X-ray photoelectron spectroscopy analysis has shown the high degree of oxidation of the $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ surface: the atomic ratio O/Ga was higher than Zn/Ga at the surface.[7] Accordingly, I investigated the impact of surface oxidation on the spatial band behavior by means of a pDOS and macroscopic averaged electrostatic potential[31] combined analysis.

Here, three oxidized surface models are considered. One is $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ with O doping the outermost layer ("PairE-5 & O-0"). The second is $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ with Zn doping the deeper ninth layer ("PairE-5 & Zn-9"), whose surface is relatively O-rich compared with the bulk region. The third is characterized by the presence of both surface O and bulk Zn dopants ("PairE-5 & O-0 & Zn-9"). Their structures, pDOSs, and macroscopic averaged electrostatic potentials are shown in Fig. 3-4.

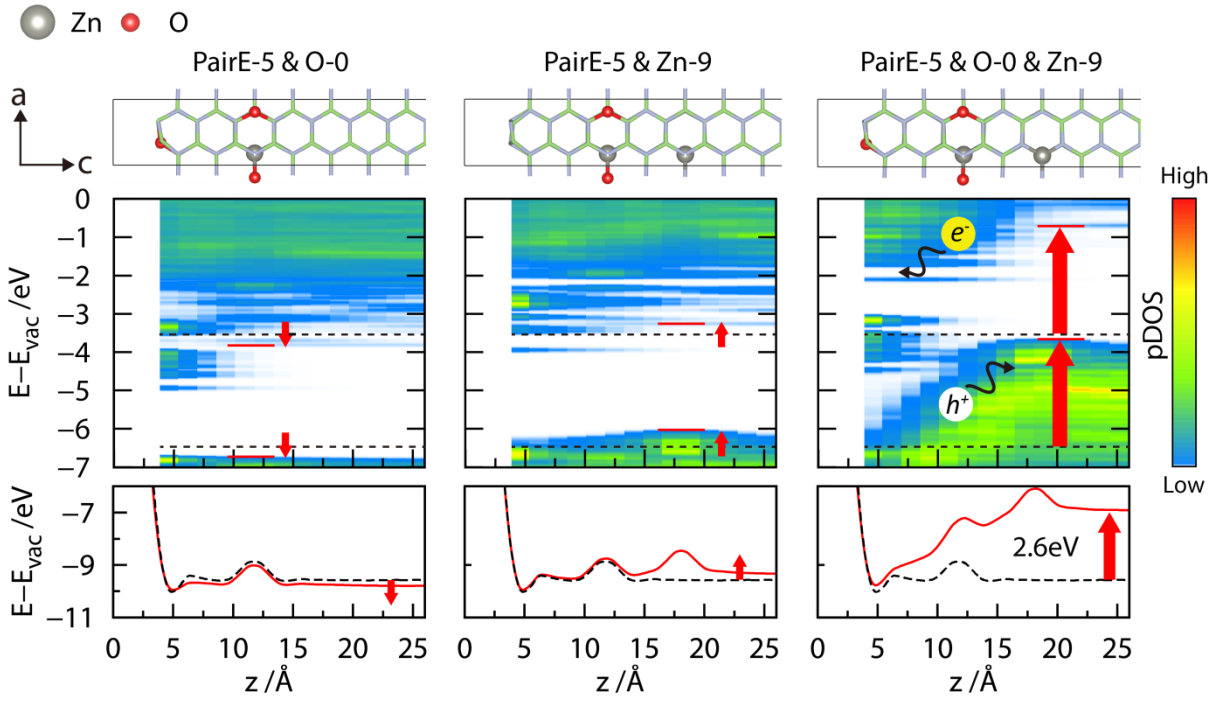


Figure 3-4. Top panel: Structures of PairE-5 & O-0 (left), PairE-5 & Zn-9 (center), and PairE-5 & O-0 & Zn-9 (right) shown from the *b*-axis direction. Middle panel: pDOS plots along the plane-normal direction. Red lines are the band-edge positions. The black dashed line is that of PairE-5. Bottom panel: Macroscopic averaged electrostatic potentials (red lines). The black dashed line is that of PairE-5.

PairE-5 & O-0 shows band-edge lowering of both the CBM and VBM compared with PairE-5. Similar energy shifts of the two band edges result in little change of the bandgap. These shifts are due to the lowered electrostatic potential in the bulk region, as shown in the bottom panel. Such a potential drop is associated with the donor character of O in GaN. To maintain the octet rule, donor O releases one valence electron, becoming slightly positive

compared with N. My Bader analysis shows $-1.21e$ on O and $-1.45e$ on N in the bulk GaN, and thus δ_{O}^{+} is $+0.24e$. Such positive character reduces the surface dipole, causing a drop of the electrostatic potential in the bulk region.

On the other hand, PairE-5 & Zn-9 model shows raising of both the CBM and VBM, and correspondingly the electrostatic potential rises. This case is reversed compared with PairE-5 & O-0; Zn is more negative than Ga because it acts as an acceptor in GaN. The Bader charge of Zn is $+0.99e$ while for Ga it is $+1.45e$, and thus δ_{Zn}^{-} is $-0.46e$. Such negative character enhances electron–electron repulsion, leading to the rise of the potential and the band edges. Here a further VBM upward shift is also observed due to the p–d repulsion on the Zn-N bonds.

Surprisingly, PairE-5 & O-0 & Zn-9 shows a quite different behavior from the other two models. The band edges and electrostatic potential are largely pushed up in the bulk region as a consequence of the downward band bending at the surface. The bending observed between plane 0 including O and plane 9 including Zn enhances *e–h* separation, i.e., electrons and holes diffuse to the surface and bulk directions, respectively. Such noticeable bending derives from the electric field caused by the positive O-rich layer ($\delta_{\mathrm{O}}^{+} = +0.25e$) and negative Zn-rich layer ($\delta_{\mathrm{Zn}}^{-} = -0.49e$). In other words, the donor and acceptor layers form local *p-n*

junction. The potential difference between bulk and surface, ΔV, is represented by the Helmholtz equation[57–59] as

$$\Delta V = \frac{12\pi\mu N_{\mathrm{dipole}}}{\varepsilon} \tag{3.1}$$

where μ, N_{dipole}, and ε are the dipole moment (Debye), surface density of dipoles (dipoles/Å^2), and dielectric constant of the medium, respectively. Assuming $\mu = \delta_{\mathrm{O}}^{+} d_{\mathrm{O-0,Zn-9}}$, where $d_{\mathrm{O-0,Zn-9}}$ is the O-0–Zn-9 distance along the z-direction (14.16 Å), and using the experimental value of 9.5 for $\varepsilon_{\mathrm{GaN}}^{\perp}$,[60] I obtain ΔV = 2.48 V. This value agrees very well with the calculated difference of electrostatic potentials from ones of PairE-5 (2.6 eV as shown in Fig. 3-4). The band bending we observed is quite large because of the high N_{dipole} associated with the limited size of my model. Although the presence of such a downward bending is not confirmed in real systems, my following discussion about photocatalytic properties holds regardless of the band bending extent.

3-3-4. Discussion: carrier dynamics in non-equilibrium condition

Experimentally, it is reported that using silver nitrate, which is a sacrificial electron acceptor, the quantum efficiency for O_2 evolution becomes markedly higher than that for overall water splitting.[6,61] It is also reported only loading H_2 evolution cocatalysts such as Rh–Cr mixed oxide enhance both H_2 and O_2 evolution, while no activity is shown by loading just an O_2 evolution cocatalyst such as Mn_3O_4.[62,63] These findings indicate that, in order to initiate OER, hole accumulation prior electron consumption is more crucial than the presence of oxidation reaction sites. My result shown in the subsection 3-3-2 supports such a hypothesis. The reason that high mandatory hole accumulation is that they are trapped on the valence states localized around the Zn-N bonds. In other words, pushing the quasi-Fermi level down allows the holes to escape from such trap states.

Since powdered semiconductors in a real photocatalytic system have contacts with both a cocatalyst metal and aqueous solution, the picture of band bending in photocatalysis is not perfectly described in my calculations. However, although previous photoelectrochemical measurements showed $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ was an n-type semiconductor,[8,61] expected upward band bending due to the electron transfer to aqueous solution has not been confirmed in the powders. Indeed, it is believed that such upward bending decreases in small particles because of the low donor concentration and the size effect.[64,65] On the other hand, the bending we found is not due to the the electron transfer at an interface. It arises regardless of the particle

sizes. Such bending is peculiar to co-doped systems including both donor and acceptor dopants, thus this band bending could be one possible explanation why $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ has higher photocatalytic activity compared with pure GaN and ZnO[5].

The bending we found is opposite to the upward bending usually occurred at the interface with an n-type semiconductor and liquid. As far as I know there is no direct measurement of band bending of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$, thus it is still controversial issue. If I assume that $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ has downward bending, it enhances electron ejection out of the surface, and hence the H_2 evolution is achieved with a suitable reaction site. In contrast, the OER is still a difficult process even with an O_2 evolution cocatalyst because the bending prohibits hole diffusion toward the surface. Therefore, in addition to the finding of a low hole mobility associated with the localized VBM nature, this could explain the experimentally reported difficulty to initiate the OER.

3-4. Bandgap and photoabsorption properties of GaN-rich and ZnO-rich alloys

3-4-1. Computational details

In Sec 3-3, we found that the number and the orientation of Zn-N bonds are correlated with the VBM positions in GaN-rich compounds. In this section, I will confirm this relation by means of GW approximation giving accurate bandgaps for both GaN-rich (x = 0.125) and ZnO-rich (x = 0.875) cases. Here bulk models are employed.

I systematically investigate tens of solid solution models by examining the local atomic configuration. Figure 3-5 shows the ten configurations of ZnO segment embedded in the GaN-rich materials (x = 0.125). The morphology generated here can be classified into three groups: Group A "*impurity*" models, Group B "*anion-cation pair*" models, and Group C "*tetramer cluster*" models. Group A and B are (2×2×1) supercells of 4-atom wurtzite primitive cell (16 atoms) and group C is (2×2×2) ones (32 atoms).

In Group A Zn and O atoms are not neighbored while they are neighbored in Group B. A1 and B1 have the Zn and O atoms in the same *ab*-plane while they lie in different planes in A2 and B2. In Group C tetramer clusters consisting of two Zn and two O atoms are formed. The six models have different tetramer conformations. I also generate ten ZnO-rich material (x = 0.875) by replacing Ga, N, Zn, and O atoms with Zn, O, Ga, N atoms, respectively.

The amount of non-isovalent bonds (Zn-N and Ga-O) is associated with the size of domains since these bonds lies on the interfaces of domains of GaN and ZnO, i.e., larger domain sizes causes more non-isovalent bonds. Therefore, each group in my models has the same number of non-isovalent bonds: group A, B, and C have both Zn-N and Ga-O bonds of 12.5, 9.375, and 7.8125 % respectively.

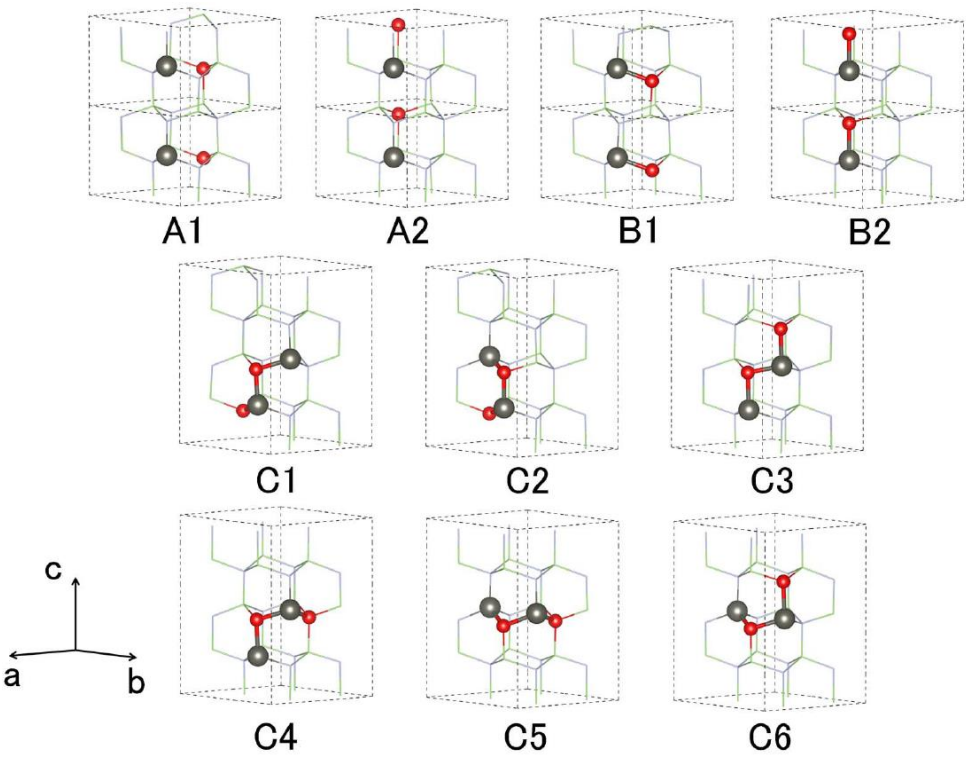


Figure 3-5. Structures of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ models in GaN-rich case (x = 0.125). Gray and red atoms represent Zn and O. To exaggerate the conformations of the ZnO domains, Ga and N are represented as thin green and blue lines. Group A and B have (2x2x1) super cells and group C has (2x2x2) ones.

My calculations are divided into two main sections: one concerns the electronic structures and the other the optical properties. For the calculations of electronic structures, i.e., DFT and GW approximation, I employed PAW scheme using VASP code.[55,66] The cutoff energy of the plane wave basis set in DFT calculations is set to 500 eV. $10/N_x \times 10/N_y \times 10/N_z$ k-grid is sampled by Monkhorst-pack scheme for $N_x \times N_y \times N_z$ super cells. The 3d-electrons are treated as valence electrons for both Ga and Zn. In GW calculations Monkhorst-pack sampling is reduced to $6/N_x \times 6/N_y \times 6/N_z$. The cutoff energy for the response function is set to 90 eV. 200 empty bands and 48 frequency points on real axis are used.

To calculate the optical properties, the independent-particle (IP) approximation is a computationally feasible way but the lack of electron-hole interaction is inadequate to compare the theoretical results with experimental absorption spectra.[20,33,37] Therefore complex dielectric functions are obtained by solving BSE with the YAMBO code.[67] KS orbitals for YAMBO are derived from DFT calculations by PWscf,[68] which agree well with the VASP calculated ones. The cutoff energies of exchange and screened Coulomb interaction are set to 40 and 3 Rydberg respectively. For comparison with experimental results, a sufficiently dense k-point is required to get smooth optical spectra. Therefore I employ the double-grid method[69] which ensures converged optical spectra with moderate computational costs.

3-4-2. Bandgaps and formation energies

The ratio of Zn-N bonds, lattice parameters, bandgaps (E_g), and formation energies (E_{form}) of all alloy models are summarized in Tab. 3-2. Their structures are optimized at LDA level. Both *a* and *c* lattice constants are larger than pure GaN and ZnO showing the deviation from Vegard's law. This is consistent with previous reports.[18,27] The lattice parameters clearly depend on the ZnO/GaN ratio, whereas the impact of the local atomic configuration is almost negligible.

As mentioned in Sec 3-2, bandgaps of GaN and ZnO are reproduced by PBE+G_0W_0 and HSE06+G_0W_0, respectively. Therefore in this work the bandgaps of GaN-rich and ZnO-rich alloys are calculated at PBE+G_0W_0 and HSE06+G_0W_0 levels. All the systems have direct bandgaps on Γ-point. E_{form} are calculated at PBE and HSE06 levels for GaN-rich and ZnO-rich models, respectively, as the total energy differences from the pure constituents

$$E_{\mathrm{form}}(x) = E_{\mathrm{alloy}} - ((1-x)E_{\mathrm{GaN}} + xE_{\mathrm{ZnO}}). \qquad (3.2)$$

E_g and E_{form} are remarkably changed in ten different structures at both GaN-rich and ZnO-rich models. Comparing model A2 with model C2, for instance, the differences are 0.78 eV and 0.90 eV in GaN-rich and ZnO-rich cases, respectively. Such large variation even at same atomic ratio has also been reported by a previous paper based on B3LYP calculations.[21] It clearly indicates the relevant role played by the atomic configuration as shown in Sec 3-3.

Table 3-2. Bandgaps, formation energies, and lattice constants of all alloy models

					Lattice constant /Å	
		Zn-N bonds /%	E_g /eV	E_{form} /meV atom^{-1}	a	c
GaN-rich	A1	12.5	2.48	37.3	3.17	5.15
	A2		2.29	36.9	3.17	5.15
	B1	9.375	2.81	5.9	3.16	5.15
	B2		2.67	13.2	3.16	5.15
	C1	7.8125	2.99	-0.2	3.16	5.15
	C2		3.07	-0.7	3.16	5.15
	C3		2.49	8.0	3.16	5.15
	C4		2.61	5.9	3.17	5.15
	C5		2.94	3.8	3.17	5.15
	C6		2.64	7.5	3.16	5.15
ZnO-rich	A1	12.5	2.25	61.4	3.20	5.17
	A2		2.05	59.8	3.20	5.17
	B1	9.375	2.73	22.5	3.20	5.16
	B2		2.58	30.7	3.19	5.17
	C1	7.8125	2.91	16.0	3.19	5.16
	C2		2.95	15.5	3.19	5.17
	C3		2.43	24.4	3.20	5.15
	C4		2.31	21.8	3.20	5.15
	C5		2.62	21.7	3.20	5.17
	C6		2.64	24.7	3.20	5.17

The impact of atomic configurations on both E_{form} and E_g is explained by the amount of non-isovalent bonds Ga-O and Zn-N; a lot of non-isovalent bonds result in higher formation energies and narrower bandgaps. The relation between bandgaps and Zn-N bonds is consistent with the results in Sec 3-3: an increased number of Zn-N bonds causes VBM upward shift and bandgap lowering. Correlation with E_{form} is also clear consequence of the fact that the mismatch of the valence electrons causes instability of the covalent bonds. Such relationship also holds for the case of another non-isovalent alloy $(GaAs)_{1-x}(Ge_2)_x$.[70,71]

For further discussing my results, E_g are plotted with the E_{form} in Fig.3-6. Dashed and solid lines are fitted by least squares method with all of GaN-rich and ZnO-rich models, respectively. The clear negative trend indicates the bandgaps are related to the stability of valence electrons, which confirms the results in subsection 3-3-2 that the impact of atomic configuration is mainly from the VBM positions. The trends in GaN-rich and ZnO-rich ones are similar. Hence, in both GaN-rich and ZnO-rich cases, the bandgap reduction is explained in terms of the number and the orientation of Zn-N bonds.

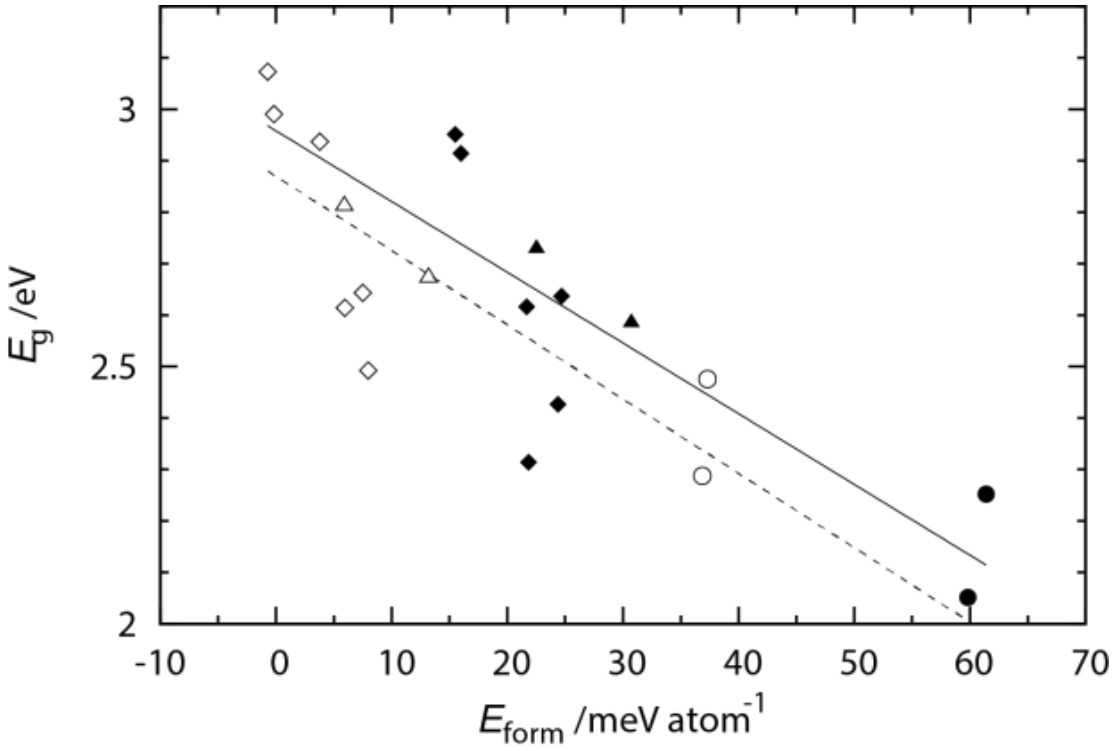


Figure 3-6. Bandgaps plotted with the formation energies of GaN-rich (empty) and ZnO-rich (filled) models. Circles, triangles, diamonds are group A, B, and C models. Dashed and solid lines are fitted with GaN-rich and ZnO-rich models respectively.

Fig. 3-7 shows VBM charge densities $|\varphi_{\mathrm{VBM}}(\mathbf{r})|^2$ of GaN-rich group C models characterized by the Zn_2O_2 tetramer cluster presence. Zn and O atoms and Zn-N and Zn-O bonds are shown while Ga and N atoms and Ga-N and Ga-O bonds are hidden for visual aid. We observe that the resulting charge densities are localized on N atoms bonded with Zn. This character confirms the picture that VBMs of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ are derived from the anti-bonding orbital of N 2p and Zn 3d.[18–21]

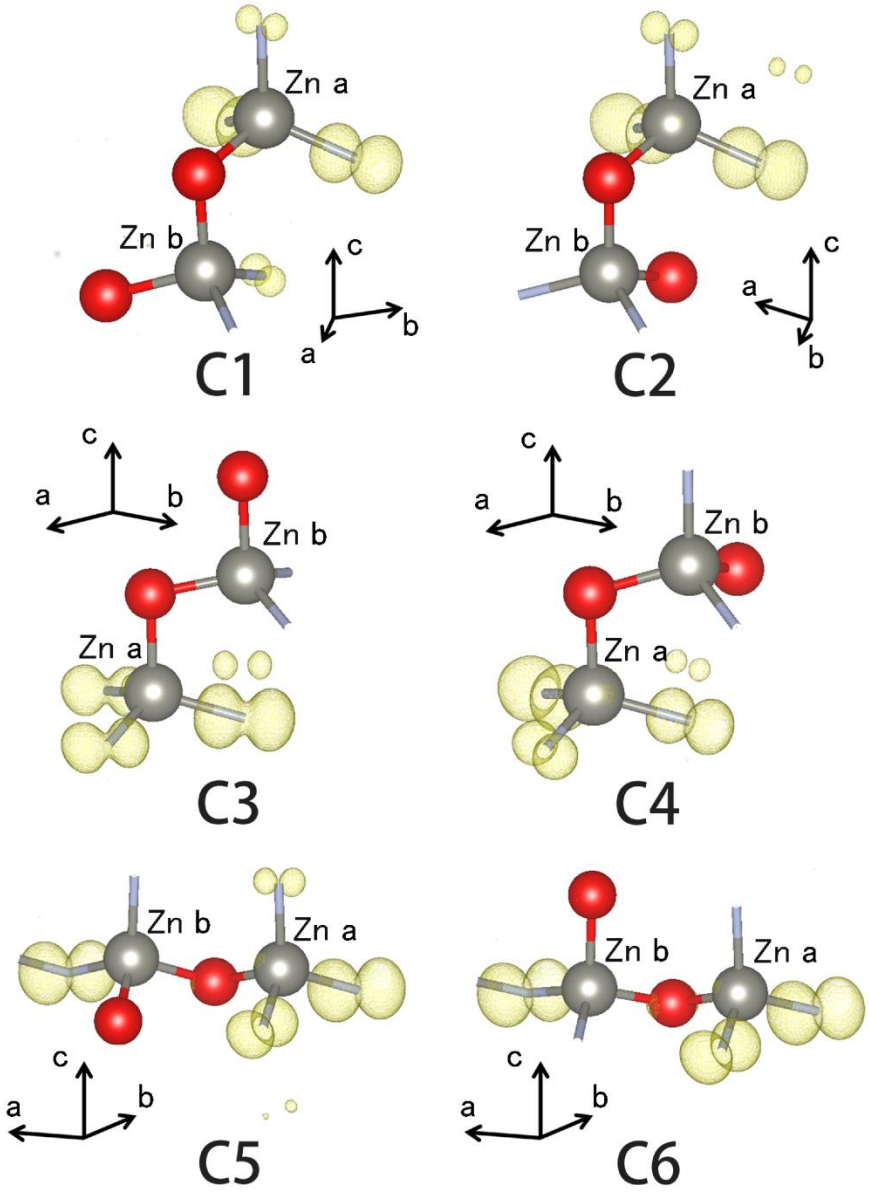


Figure 3-7. Charge densities of VBMs of GaN-rich models with tetramer Zn_2O_2 cluster. Zn (O) atoms are shown gray (red) and all Ga and N atoms are hidden for the visual aid. The Zn-O and Zn-N bonds are shown.

The two zinc atoms in Zn_2O_2 tetramer clusters are not equivalent; indeed, one is bound to three nitrogen atoms and one oxygen atom (ZnN_3O) and the other bound to two nitrogen atoms and two oxygen atoms (ZnN_2O_2). Hereafter I will call the former "Zn-*a*" and the latter "Zn-*b*". Fig. 3-7 clearly shows that the charge densities are localized on N atoms bound to Zn-*a*. Such localization stems from the fact that the uppermost valence orbital originates from the most electronically coupled Zn and N atom orbitals. In C5 and C6, the charge density also resides in the N atom nearest neighbor of Zn-*b*. This is because those N atoms are also bound to the Zn-*a* in the next periodic cell.

This trend of the VBM localization is supposed to cause the hole mobility reduction, as discussed in Sec 3-3. In ZnO-rich models, the charge distribution of the VBM is different from the GaN-rich case since the contribution of N 2p and Zn 3d orbitals is not equivalent. However, similar trend of charge localization is observed.

3-4-3. Optical properties

Bandgaps comparable with experiment are obtained by GW approximation in the previous section. However, in order to compare the experimental photoabsorption properties, direct calculations of absorption spectra, not just absorption edges, are preferable. It is of paramount relevance in the case of GaN:ZnO because the experimental absorption spectra shows very broad peak leading to unclear absorption edges.

Complex dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$ (Eq. (2.106)) including electron-hole interaction was computed by solving BSE. The calculations were performed on top of the KS orbitals by LDA+U calculations with scissors shift to match the bandgaps from GW approximation shown in Table 3-2. Fig. 3-8 shows the absorption coefficient,

$$\alpha(\omega) = \sqrt{2}\omega/c(|\varepsilon(\omega)| - \varepsilon_1(\omega))^{1/2}, \qquad (3.3)$$

comparing with the experimental absorbance from UV-Vis DRS.[14,16,17,22] Double grid method for k-point sampling enables to obtain converged absorption spectra, thus direct comparison with experimental spectra is allowed .

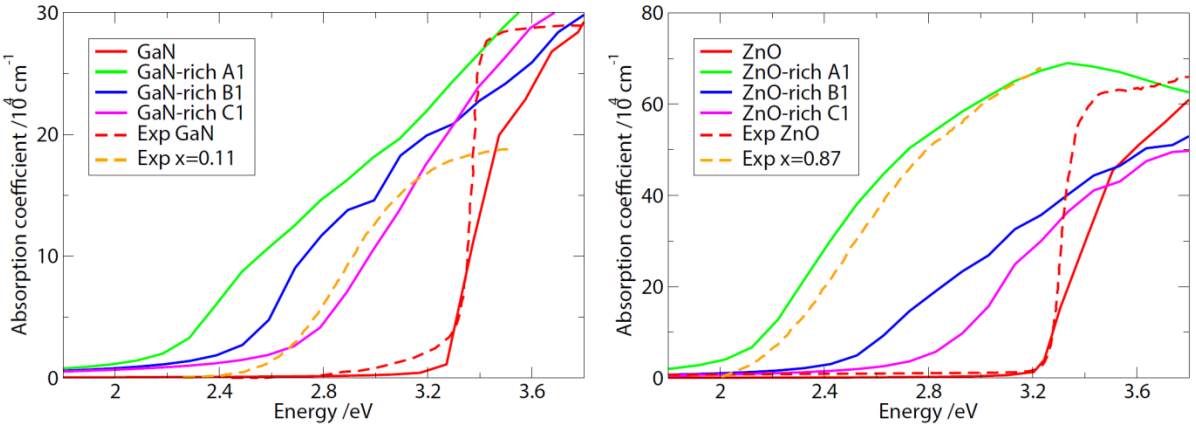


Figure 3-8. Absorption coefficients calculated by BSE with double grid method. Left panel: pure GaN and GaN-rich A1, B1, and C1 alloys, Right panel: pure ZnO and ZnO-rich A1, B1, and C1 alloys. Experimental results of pure GaN and ZnO are from Ref [14] and Ref [16]. $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ of x = 0.11 and x = 0.87 from Ref [22] and Ref [17] respectively.

My results for pure, parental GaN and ZnO show excellent agreement in the absorption edge positions with the experiment due to the inclusion of excitonic effects. The sharpness of the edges is, however, not completely reproduced as shown in other theoretical papers.[54,72] This can be ascribable to the absence of band stretching correction on the band structures. However, such problem is not pronounced in the case of solid solution models because their band dispersion is quite small due to the localization of the wave function, as shown in Fig. 3-7.

For comparison with the experimental absorbance, A1, B1, and C1 models are chosen for the optical property calculations. All of solid-solution models show gradual absorption edges, deriving from the small overlap of wave functions of the VB and the CB, i.e., the VBs of

these alloys are mainly localized on the N atoms bonding with Zn atoms while the CBs are delocalized in overall the crystals, as shown in Fig. 3-3 and 3-7. In both GaN-rich and ZnO-rich cases, the model A1 has the strongest intensity as the consequence of the highest amount of Zn-N bonds that maximizes the overlap of the VB and the CB.

All my models are still regarded as impurity-like system introducing deep N 2p-Zn 3d coupled acceptor levels with the compensating donors (O in GaN-rich case and Ga in ZnO-rich case, respectively). Such a picture is consistent with both the p-d repulsion[18–21] and the impurity-like picture[22] previously argued.

In three different atomic configurations, GaN-rich C1 model (Zn-N bonds of 7.8125 %) and ZnO-rich A1 model (Zn-N bonds of 12.5 %) exhibit the best agreement with experimental spectra. This result suggests the possibility that different atomic configurations are achieved in GaN-rich and ZnO-rich solid solutions: there are more Zn-N bonds in ZnO-rich case than that in GaN-rich case. In other words, we can say that Zn and O atoms partially clusterize in GaN matrix in GaN-rich compounds while, in ZnO-rich compounds, Ga and N atoms tend to be separately embedded in ZnO matrix. The origin of such different atomic configuration is ascribed to the experimentally available methodology of synthesis and will be discussed in the next subsection.

3-4-4. Discussion: synthetic method and optical gap

Here, I will discuss the origin of the different atomic configuration in GaN-rich and ZnO-rich cases as suggested in Fig. 3-8.

Maeda's traditional method to synthesize GaN-rich alloys, which was picked up for the comparison with my theoretical absorbance (x = 0.11 in Fig. 3-8), consists in the nitridation of mixtures of β-Ga_2O_3 and ZnO at T = 1123 K for 10 h.[22] In situ time-resolved XRD study shows that the wurtzite $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ is formed after the formation of the spinel $ZnGa_2O_4$ intermediated phase,[73] indicating that the important step to form final solid solution is the migration of nitrogen atoms into metal oxide. Accordingly, we can give a rational explanation from the stochastic point of view: migrating N atoms are likely to bond with Ga since $ZnGa_2O_4$ consists in a larger amount of Ga than Zn. Furthermore, Zn-O pairing is also thermodynamically favorable in the wurtzite lattice as reported by MC simulations.[26,27] Hence ZnO clusters are preferably formed, which leads to the subsequent reduction of Zn-N bond. This is consistent with my C1 model which reproduced the experiment in Fig. 3-8.

On the other hand, the synthesis method of ZnO-rich case by Lee et al. (experiment of x = 0.87 in Fig. 3-8) is the nitridation of mixtures of $ZnGa_2O_4$ and ZnO nanocrystals under T = 923 K for 10 h.[17] In this case, in contrast with Maeda's method, the precursor includes much Zn, thus migrating N atoms are likely to form bonds with Zn even though Ga-N bond is thermodynamically preferred. In order to achieve Ga-N pairing formation in the solid

solutions, sufficiently high kinetic energies for atoms to diffuse freely and overcome the activation barriers are required. However, in order to prevent Zn evaporation, the temperature of this method is much lower than the traditional method, thus the kinetic energies might be insufficient. Thus the Ga and N atoms could be distributed independently as in my impurity-like A1 model which reproduced the experiment in Fig. 3-8. Other ways to synthesize ZnO-rich alloys by solution combustion method[15] or nitridation of LDHs[16] also have largely red-shifted absorption edges (~2.2 eV at x = 0.88 and 2.37 eV at x = 0.81, respectively), which are explained in a similar manner because the precursors in these reactions also include much more Zn than Ga. Actually, certain amount of Zn-N bonds was confirmed from the Raman spectra and secondary ion mass spectrometry in the samples by solution combustion method.[15] In the case of LDHs nitridation the temperature is relatively high (T = 1073 K), but in this case the incompleteness of the equilibration can be ascribed to the short reaction time, just 30 min.

Previous experimental study has associated the gap narrowing to the poor crystallinity in low-temperature synthetic regime.[74] My finding seems related to such result because a large number of Zn-N bonds also causes structural disorder, thus the crystallinity of the sample result poor. I have not considered the impact of non-stoichiometry, which also worsens the crystallinity. It will be addressed in further future investigations.

Finally, I comment on the mechanism of bandgap reduction at intermediate concentrations (x ~ 0.5). I believe the mechanism of bandgap reduction is different in dilute and intermediate concentrations. Previous soft XAS study[23] for GaN-rich solid solutions indicates that a sizable fraction of ZnO remains in GaN host at x = 0.18 and 0.42, but not at x = 0.06 (Fig 5 of Ref [23]), which indicates that Zn and O start to aggregate between x = 0.06 and 0.18. At low ZnO concentration before the aggregation starts, amount of Zn-N bonds increases as function of x. However, after the aggregation starts for intermediate concentrations, the Zn-O network prevents the amount of Zn-N bond to increase, therefore the bandgap reduction due to VBM raising is inhibited. In such intermediate concentration regions, the bandgap change with x can be mainly attributed to the CB repulsion as stressed by McDermott et al.[23]

3-5. Conclusion

In this chapter, I have investigated the mechanism of visible light response in $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ by means of first-principles calculations focusing, in particular, on the impact of local atomic configuration.

In Sec 3-2, I demonstrated test of bandgap calculations by LDA, LDA+U, and GW approximation for pure GaN and ZnO. As widely reported, LDA predicted underestimated bandgaps compared with the experiment. LDA+U corrects the positions of cation d bands, resulting in improved bandgaps although still not satisfactorily in the case of ZnO. I confirmed that PBE+G_0W_0 and HSE+G_0W_0 provide good agreement for GaN and ZnO, respectively. The problem of GW approximation is the heavy computational cost, hence I have opted for feasible LDA+U for slab models with more than one hundred atoms for the GaN-rich case. GW approximation is employed to calculate accurate bandgap of smaller systems.

In Sec 3-3, LDA+U calculations for GaN-rich slab models (x=0.0345) with $(11\bar{2}0)$ facets were reported. Calculations revealed that the bandgap lowering by chemical alloying stems from both VBM raising and CBM lowering. Only the position of VBMs strongly depends on the atomic configuration, that is, the number and orientation of Zn-N bonds, while that of CBMs results not influenced. I also found that the hole mobility is reduced since the VBM orbital is localized around the Zn-N bonds. Furthermore, I observed downward band bending

at the O-rich GaN:ZnO surface due to the electric field induced by opposite charges of donor O and acceptor Zn (local *p*-*n* junction). These results seem related to the fact that GaN:ZnO requires high hole accumulation for water-splitting reaction.

In Sec 3-4, by means of the GW approximation, I have investigated several bulk models of both GaN-rich and ZnO-rich concentrations (x = 0.125 and 0.875). The trend reported in Sec 3-3, i.e., an increased number of Zn-N bonds that induces bandgap narrowing, is confirmed also in the case of ZnO-rich alloys. The comparison of absorption spectra computed by BSE with experiments indicates that the atomic configuration is different in GaN-rich and ZnO-rich cases; ZnO-rich solid solutions have larger number of Zn-N bonds compared with GaN-rich ones. Such difference can be ascribed to the different methodologies of synthesis.

Bibliography

(1) Fujishima, A.; Honda, K. *Nature* **1972**, *238*, 37.

(2) Kudo, A.; Miseki, Y. *Chem. Soc. Rev.* **2009**, *38*, 253.

(3) Moriya, Y.; Takata, T.; Domen, K. *Coord. Chem. Rev.* **2013**, *257*, 1957.

(4) Maeda, K.; Domen, K. *Chem. Mater.* **2010**, *22*, 612.

(5) Maeda, K.; Takata, T.; Hara, M.; Saito, N.; Inoue, Y.; Kobayashi, H.; Domen, K. *J. Am. Chem. Soc.* **2005**, *127*, 8286.

(6) Maeda, K.; Teramura, K.; Lu, D.; Takata, T.; Saito, N.; Inoue, Y.; Domen, K. *Nature* **2006**, *440*, 295.

(7) Maeda, K.; Teramura, K.; Takata, T.; Hara, M.; Saito, N.; Toda, K.; Inoue, Y.; Kobayashi, H.; Domen, K. *J. Phys. Chem. B* **2005**, *109*, 20504.

(8) Hashiguchi, H.; Maeda, K.; Abe, R.; Ishikawa, A.; Kubota, J.; Domen, K. *Bull. Chem. Soc. Jpn.* **2009**, *82*, 401.

(9) Maeda, K.; Teramura, K.; Domen, K. *J. Catal.* **2008**, *254*, 198.

(10) Ohno, T.; Bai, L.; Hisatomi, T.; Maeda, K.; Domen, K. *J. Am. Chem. Soc.* **2012**, *134*, 8254.

(11) Brook, S.; York, N. **2000**, *85*, 543.

(12) Stevanović, V.; Lany, S.; Ginley, D. S.; Tumas, W.; Zunger, A. *Phys. Chem. Chem. Phys.* **2014**, *16*, 3706.

(13) Maeda, K.; Teramura, K.; Saito, N.; Inoue, Y.; Domen, K. *Bull. Chem. Soc. Jpn.* **2007**, *80*, 1004.

(14) Chen, H.; Wang, L.; Bai, J.; Hanson, J. C.; Warren, J. B.; Muckerman, J. T.; Fujita, E.; Rodriguez, J. A. *J. Phys. Chem. C* **2010**, *114*, 1809.

(15) Mapa, M.; Thushara, K. S.; Saha, B.; Chakraborty, P.; Janet, C. M.; Viswanath, R. P.; Madhavan Nair, C.; Murty, K. V. G. K.; Gopinath, C. S. *Chem. Mater.* **2009**, *21*, 2973.

(16) Wang, J.; Huang, B.; Wang, Z.; Wang, P.; Cheng, H.; Zheng, Z.; Qin, X.; Zhang, X.; Dai, Y.; Whangbo, M.-H. *J. Mater. Chem.* **2011**, *21*, 4562.

(17) Lee, K.; Tienes, B. M.; Wilker, M. B.; Schnitzenbaumer, K. J.; Dukovic, G. *Nano Lett.* **2012**, *12*, 3268.

(18) Jensen, L. L.; Muckerman, J. T.; Newton, M. D. *J. Phys. Chem. C* **2008**, *112*, 3439.

(19) Wei, W.; Dai, Y.; Yang, K.; Guo, M.; Huang, B. *J. Phys. Chem. C* **2008**, *112*, 15915.

(20) Huda, M.; Yan, Y.; Wei, S.-H.; Al-Jassim, M. *Phys. Rev. B* **2008**, *78*, 195204.

(21) Di Valentin, C. *J. Phys. Chem. C* **2010**, *114*, 7054.

(22) Yoshida, M.; Hirai, T.; Maeda, K.; Saito, N.; Kubota, J.; Kobayashi, H.; Inoue, Y.; Domen, K. *J. Phys. Chem. C* **2010**, *114*, 15510.

(23) McDermott, E. J.; Kurmaev, E. Z.; Boyko, T. D.; Finkelstein, L. D.; Green, R. J.; Maeda, K.; Domen, K.; Moewes, A. *J. Phys. Chem. C* **2012**, *116*, 7694.

(24) Wang, L.; Zunger, A. *Phys. Rev. B* **2003**, *68*, 125211.

(25) Glicksman, M.; Kraeft, W. D. *Solid. State. Electron.* **1985**, *28*, 151.

(26) Wang, S.; Wang, L.-W. *Phys. Rev. Lett.* **2010**, *104*, 065501.

(27) Li, L.; Muckerman, J. T.; Hybertsen, M. S.; Allen, P. B. *Phys. Rev. B* **2011**, *83*, 134202.

(28) Toroker, M. C.; Kanan, D. K.; Alidoust, N.; Isseroff, L. Y.; Liao, P.; Carter, E. a. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16644.

(29) Payne, M. C.; Arias, T. A.; Joannopoulos, J. D. *Rev. Mod. Phys.* **1992**, *64*, 1045.

(30) Li, Y.; Li, Y.-L.; Araujo, C. M.; Luo, W.; Ahuja, R. *Catal. Sci. Technol.* **2013**, *3*, 2214.

(31) Colombo, L.; Resta, R.; Baroni, S. *Phys. Rev. B* **1991**, *44*, 5572.

(32) Du, Y. a; Chen, Y.-W.; Kuo, J.-L. *Phys. Chem. Chem. Phys.* **2013**, *15*, 19807.

(33) Wang, Z.; Zhao, M.; Wang, X.; Xi, Y.; He, X.; Liu, X.; Yan, S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15693.

(34) Wei, S.-H.; Zunger, A. *Phys. Rev. B* **1988**, *37*, 8958.

(35) Dudarev, S. L.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P. *Phys. Rev. B* **1998**, *57*, 1505.

(36) Onida, G.; Reining, L.; Rubio, A. *Rev. Mod. Phys.* **2002**, *74*, 601.

(37) Dou, M.; Baldissera, G.; Persson, C. *J. Cryst. Growth* **2012**, *350*, 17.

(38) Vurgaftman, I.; Meyer, J. R.; Ram-Mohan, L. R. *J. Appl. Phys.* **2001**, *89*, 5815.

(39) Desgreniers, S. *Phys. Rev. B* **1998**, *58*, 14102.

(40) Shishkin, M.; Kresse, G. *Phys. Rev. B* **2007**, *75*, 235102.

(41) Fuchs, F.; Furthmüller, J.; Bechstedt, F.; Shishkin, M.; Kresse, G. *Phys. Rev. B* **2007**, *76*, 115109.

(42) Lambrecht, W.; Segall, B.; Strite, S.; Martin, G.; Agarwal, A.; Morkoç, H.; Rockett, A. *Phys. Rev. B* **1994**, *50*, 14155.

(43) Mang, A.; Reimann, K.; Rübenacke, S. *Solid State Commun.* **1995**, *94*, 251.

(44) Girard, R. T.; Tjernberg, O.; Chiaia, G.; Söderholm, S.; Karlsson, U. O.; Wigren, C.; Nylén, H.; Lindau, I. *Surf. Sci.* **1997**, *373*, 409.

(45) Preston, A. R. H.; Ruck, B. J.; Piper, L. F. J.; DeMasi, A.; Smith, K. E.; Schleife, A.; Fuchs, F.; Bechstedt, F.; Chai, J.; Durbin, S. M. *Phys. Rev. B* **2008**, *78*, 155114.

(46) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048.

(47) Perdew, J.; Levy, M. *Phys. Rev. B* **1997**, *56*, 16021.

(48) Perdew, J.; Levy, M. *Phys. Rev. Lett.* **1983**, *51*, 1884.

(49) Sham, L.; Schlüter, M. *Phys. Rev. Lett.* **1983**, *51*, 1888.

(50) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(51) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207.

(52) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2006**, *124*, 219906.

(53) De Carvalho, L. C.; Schleife, A.; Bechstedt, F. *Phys. Rev. B* **2011**, *84*, 195105.

(54) Schleife, A.; Rödl, C.; Fuchs, F.; Furthmüller, J.; Bechstedt, F. *Phys. Rev. B* **2009**, *80*, 035112.

(55) Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996**, *6*, 15.

(56) Van de Walle, C. G.; Neugebauer, J. *Nature* **2003**, *423*, 626.

(57) Martin, P.; Bornong, B. J. *J. Phys. Chem.* **1968**, *72*, 4172.

(58) Bermudez, V. M. *J. Appl. Phys.* **1996**, *80*, 1190.

(59) Li, W.-X.; Stampfl, C.; Scheffler, M. *Phys. Rev. B* **2002**, *65*, 075407.

(60) Barker, A.; Ilegems, M. *Phys. Rev. B* **1973**, *7*, 743.

(61) Maeda, K.; Hashiguchi, H.; Masuda, H.; Abe, R.; Domen, K. *J. Phys. Chem. C* **2008**, *112*, 3447.

(62) Maeda, K.; Xiong, A.; Yoshinaga, T.; Ikeda, T.; Sakamoto, N.; Hisatomi, T.; Takashima, M.; Lu, D.; Kanehara, M.; Setoyama, T.; Teranishi, T.; Domen, K. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 4096.

(63) Xiong, A.; Yoshinaga, T.; Ikeda, T.; Takashima, M.; Hisatomi, T.; Maeda, K.; Setoyama, T.; Teranishi, T.; Domen, K. *Eur. J. Inorg. Chem.* **2014**, *2014*, 767.

(64) Ohtani, B. *J. Photochem. Photobiol. C Photochem. Rev.* **2010**, *11*, 157.

(65) Zhang, Z.; Yates, J. T. *Chem. Rev.* **2012**, *112*, 5520.

(66) Shishkin, M.; Kresse, G. *Phys. Rev. B* **2006**, *74*, 035101.

(67) Marini, A.; Hogan, C.; Grüning, M.; Varsano, D. *Comput. Phys. Commun.* **2009**, *180*, 1392.

(68) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. *J. Phys. Condens. Matter* **2009**, *21*, 395502.

(69) Kammerlander, D.; Botti, S.; Marques, M. A. L.; Marini, A.; Attaccalite, C. *Phys. Rev. B* **2012**, *86*, 125203.

(70) Kawai, H.; Giorgi, G.; Yamashita, K. *Phys. status solidi* **2012**, *249*, 29.

(71) Kawai, H.; Giorgi, G.; Yamashita, K. *Chem. Lett.* **2011**, *40*, 770.

(72) Gori, P.; Rakel, M.; Cobet, C.; Richter, W.; Esser, N.; Hoffmann, A.; Del Sole, R.; Cricenti, A.; Pulci, O. *Phys. Rev. B* **2010**, *81*, 125207.

(73) Chen, H.; Wen, W.; Wang, Q.; Hanson, J. C.; Muckerman, J. T.; Fujita, E.; Frenkel, A. I.; Rodriguez, J. A. *J. Phys. Chem. C* **2009**, *113*, 3650.

(74) Ward, M. J.; Han, W.-Q.; Sham, T.-K. *J. Phys. Chem. C* **2013**, *117*, 20332.

Chapter 4.
Electron-phonon Coupling Effects on Optical Absorption and Carrier Lifetime

As described in Sec 1-2, carrier "collision" lifetime τ^{coll} is governed by scattering processes with electrons, phonons, and impurities. In this chapter, electron-phonon (e-ph) interaction is considered in the MBPT scheme. MBPT calculations provide band broadenings due to e-ph interaction, which is the inverse of τ^{coll} itself.

In Sec 4-1, theoretical background of e-ph MBPT is presented. In Sec 4-2, e-ph MBPT calculation is applied to zinc-blende GaN to demonstrate the impact of e-ph interactions on the electronic and optical properties. In Sec 4-3, e-ph MBPT is applied to investigate hot carriers in $APbI_3$, which is a promising material in perovskite solar cells.

4-1. Electron-phonon interaction in MBPT scheme

In Chapter 2, MBPT scheme for the electron-electron (e-e) and electron-hole (e-h) interactions are presented. In a similar way to the treatment of e-e interaction, we can consider e-ph interaction once we consider nuclear motions as a perturbation $\hat{V}$ in Eq (2.22). In this

scheme, the quasiparticle is so-called *polaron*. In this section the derivation of the self-energies from e-ph interaction is briefly described. More detailed derivation can be found in Ref [1].

Here we consider the total Hamiltonian of the coupled electron-nuclei system $\hat{H}$ with two interactions,

$$\hat{H} = \hat{H}_0 + \hat{H}_1 + \hat{H}_2 \tag{4.1}$$

where $\hat{H}_0$ is the electronic Hamiltonian corresponding to the case where the atoms are frozen at their equilibrium positions $\mathbf{R}_0$. Here, we employ KS-DFT to represent the electronic part;

$$\hat{H}_0 = \sum_i \left[-\frac{1}{2}\frac{\partial^2}{\partial \hat{\mathbf{r}}_i^2} + \hat{V}_{\mathrm{scf}}(\mathbf{r}_i, \mathbf{R})\big|_{\mathbf{R}=\mathbf{R_0}} \right], \tag{4.2}$$

with the electronic effective potential $\hat{V}_{\mathrm{scf}} = \hat{V}_{\mathrm{ion}} + \hat{V}_{\mathrm{H}} + \hat{V}_{\mathrm{xc}}$.

$\hat{H}_1$ and $\hat{H}_2$ represent the first and second term in the Taylor expansion of $\hat{H}_0$ when the atomic positions $\{\mathbf{R}\}$ are expanded around the equilibrium positions $\{\mathbf{R}_0\}$:

$$\hat{H}_1 = \sum_{Is\alpha} \left.\frac{\partial \hat{V}_{\mathrm{scf}}(\mathbf{r}, \mathbf{R})}{\partial R_{Is\alpha}}\right|_{\mathbf{R}=\mathbf{R_0}} \hat{u}_{Is\alpha}\,, \tag{4.3}$$

$$\hat{H}_2 = \frac{1}{2}\sum_{Is\alpha}\sum_{Js'\beta} \left.\frac{\partial^2 \hat{V}_{\mathrm{scf}}(\mathbf{r}, \mathbf{R})}{\partial R_{Is\alpha}\partial R_{Js'\beta}}\right|_{\mathbf{R}=\mathbf{R_0}} \hat{u}_{Is\alpha}\,\hat{u}_{Js'\beta}\,, \tag{4.4}$$

where $\hat{u}_{Is\alpha}$ represents lattice displacement of atom *s* in lattice *I* along with the Cartesian coordinate α. In the second quantization, the lattice displacement can be represented in the space of the canonical lattice vibrations as follows,

$$\hat{u}_{Isa} = \sum_{\mathbf{q}\lambda} \left(2N_{\mathbf{q}} M_s \omega_{\mathbf{q}\lambda}\right)^{-\frac{1}{2}} \xi_\alpha(\mathbf{q}\lambda|s) e^{i\mathbf{q}\cdot(\mathbf{R}_I+\boldsymbol{\tau}_s)} \left(b^\dagger{}_{-\mathbf{q}\lambda} + b_{\mathbf{q}\lambda}\right), \tag{4.5}$$

where $N_{\mathbf{q}}$ is the number of cells used in the simulations and M_s is the atomic mass whose position is $\boldsymbol{\tau}_s$ in the unit cell. $\omega_{\mathbf{q}\lambda}$ and $\xi_\alpha(\mathbf{q}\lambda|s)$ are phonon energy and phonon polarization vectors relative to the mode λ and transferred momentum $\mathbf{q}$. $b^\dagger{}_{-\mathbf{q}\lambda}$ and $b_{\mathbf{q}\lambda}$ are the bosonic creation and annihilation operators. Substituting Eq (4.5) into Eq (4.3), the matrix elements of $\hat{H}_1$ expanded by Bloch states can be represented as follows

$$\begin{aligned}
&\langle n\mathbf{k}|\hat{H}_1|n'\mathbf{k}'\rangle \\
&= \sum_{s\alpha}\sum_{\mathbf{q}\lambda} \left(2N_{\mathbf{q}} M_s \omega_{\mathbf{q}\lambda}\right)^{-\frac{1}{2}} \xi_\alpha(\mathbf{q}\lambda|s) e^{i\mathbf{q}\cdot\boldsymbol{\tau}_s} \\
&\qquad \times \sum_I \left\langle n\mathbf{k}\left| e^{i\mathbf{q}\cdot\mathbf{R}_I} \frac{\partial \hat{V}_{\mathrm{scf}}(\mathbf{r})}{\partial R_{Is\alpha}} \right| n'\mathbf{k}' \right\rangle \left(b^\dagger{}_{-\mathbf{q}\lambda} + b_{\mathbf{q}\lambda}\right) \\
&= \left(N_{\mathbf{q}}\right)^{-\frac{1}{2}} \sum_{\mathbf{q}\lambda} g^{\mathbf{q}\lambda}_{nn'\mathbf{k}}\, \delta_{\mathbf{k}',\mathbf{k}-\mathbf{q}} \left(b^\dagger{}_{-\mathbf{q}\lambda} + b_{\mathbf{q}\lambda}\right),
\end{aligned} \tag{4.6}$$

where $g^{\mathbf{q}\lambda}_{nn'\mathbf{k}}$ is the electron-phonon matrix element defined as

$$g^{\mathbf{q}\lambda}_{nn'\mathbf{k}} = \sum_{s\alpha} \frac{\xi_\alpha(\mathbf{q}\lambda|s)}{\left(2M_s\omega_{\mathbf{q}\lambda}\right)^{\frac{1}{2}}} e^{i\mathbf{q}\cdot\boldsymbol{\tau}_s} \langle n\mathbf{k}\left|\frac{\partial \hat{V}_{\mathrm{scf}}(\mathbf{r})}{\partial R_{s\alpha}}\right| n'\mathbf{k}-\mathbf{q}\rangle, \tag{4.7}$$

with $R_{s\alpha} = R_{Is\alpha}|_{I=0}$. In the case of second derivative, we employ the approximation to neglect the nondiagonal terms $I \neq J$ and $s \neq s'$ as

$$\frac{\partial^2 \hat{V}_{\mathrm{scf}}(\mathbf{r},\mathbf{R})}{\partial R_{Is\alpha}\partial R_{Js'\beta}} \approx \delta_{IJ}\delta_{ss'} \frac{\partial^2 \hat{V}_{\mathrm{scf}}(\mathbf{r},\mathbf{R})}{\partial R_{Is\alpha}\partial R_{Is\beta}}. \tag{4.8}$$

Then, similarly, substituting Eq (4.5) into Eq (4.4), the matrix elements of $\hat{H}_2$ is written

$$
\begin{aligned}
&\langle n\mathbf{k}|\hat{H}_2|n'\mathbf{k}'\rangle \\
&= \frac{1}{2}\sum_{s\alpha\beta}\sum_{\mathbf{q}\lambda}\sum_{\mathbf{q}'\lambda'}\left(2N_{\mathbf{q}}M_s\right)^{-1}\left(\omega_{\mathbf{q}\lambda}\omega_{\mathbf{q}'\lambda'}\right)^{-\frac{1}{2}}\xi_\alpha^*(\mathbf{q}\lambda|s)\xi_\beta(\mathbf{q}'\lambda'|s)e^{i(\mathbf{q}+\mathbf{q}')\cdot\boldsymbol{\tau}_s} \\
&\quad\times\sum_I\left\langle n\mathbf{k}\left|e^{i(\mathbf{q}+\mathbf{q}')\cdot\mathbf{R}_I}\frac{\partial^2\hat{V}_{\mathrm{scf}}(\mathbf{r})}{\partial R_{Is\alpha}\partial R_{Is\beta}}\right|n'\mathbf{k}'\right\rangle\left(b^\dagger{}_{-\mathbf{q}\lambda}+b_{\mathbf{q}\lambda}\right)\left(b^\dagger{}_{-\mathbf{q}'\lambda'}+b_{\mathbf{q}'\lambda'}\right) \\
&= \sum_{\mathbf{q}\lambda}\sum_{\mathbf{q}'\lambda'}\left(N_{\mathbf{q}}\right)^{-1}\Lambda_{nn'\mathbf{k}}^{\mathbf{q}\lambda,\mathbf{q}'\lambda'}\,\delta_{\mathbf{k}',\mathbf{k}-\mathbf{q}-\mathbf{q}'}\left(b^\dagger{}_{-\mathbf{q}\lambda}+b_{\mathbf{q}\lambda}\right)\left(b^\dagger{}_{-\mathbf{q}'\lambda'}+b_{\mathbf{q}'\lambda'}\right),
\end{aligned}
\tag{4.9}
$$

where $\Lambda_{nn\mathbf{k}}^{\mathbf{q}\lambda,-\mathbf{q}\lambda}$ is a second-order electron-phonon matrix element:

$$
\Lambda_{nn'\mathbf{k}}^{\mathbf{q}\lambda,\mathbf{q}'\lambda'} = \frac{1}{2}\sum_{s\alpha\beta}\frac{\xi_\alpha^*(\mathbf{q}\lambda|s)\xi_\beta(\mathbf{q}'\lambda'|s)}{2M_s\left(\omega_{\mathbf{q}\lambda}\omega_{\mathbf{q}'\lambda'}\right)^{\frac{1}{2}}}e^{i(\mathbf{q}+\mathbf{q}')\cdot\boldsymbol{\tau}_s}\langle n\mathbf{k}\left|\frac{\partial^2\hat{V}_{\mathrm{scf}}(\mathbf{r})}{\partial R_{s\alpha}\partial R_{s\beta}}\right|n'\mathbf{k}-\mathbf{q}-\mathbf{q}'\rangle\,. \tag{4.10}
$$

As a result, $\hat{H}_1$ and $\hat{H}_2$ are rewritten as

$$
\begin{aligned}
\hat{H}_1 &= \sum_{n\mathbf{k}}\sum_{n'\mathbf{k}'}\langle n\mathbf{k}|\hat{H}_1|n'\mathbf{k}'\rangle\, c_{n\mathbf{k}}^\dagger c_{n'\mathbf{k}'} \\
&= \left(N_{\mathbf{q}}\right)^{-\frac{1}{2}}\sum_{nn'\mathbf{k}}\sum_{\mathbf{q}\lambda} g_{nn'\mathbf{k}}^{\mathbf{q}\lambda}\, c_{n\mathbf{k}}^\dagger c_{n'\mathbf{k}-\mathbf{q}}\left(b^\dagger{}_{-\mathbf{q}\lambda}+b_{\mathbf{q}\lambda}\right),
\end{aligned}
\tag{4.11}
$$

and

$$
\begin{aligned}
\hat{H}_2 &= \sum_{n\mathbf{k}}\sum_{n'\mathbf{k}'}\langle n\mathbf{k}|\hat{H}_2|n'\mathbf{k}'\rangle\, c_{n\mathbf{k}}^\dagger c_{n'\mathbf{k}'} \\
&= \left(N_{\mathbf{q}}\right)^{-1}\sum_{nn'\mathbf{k}}\sum_{\mathbf{q}\lambda}\sum_{\mathbf{q}'\lambda'}\Lambda_{nn'\mathbf{k}}^{\mathbf{q}\lambda,\mathbf{q}'\lambda'}\, c_{n\mathbf{k}}^\dagger c_{n'\mathbf{k}-\mathbf{q}-\mathbf{q}'}\left(b^\dagger{}_{-\mathbf{q}\lambda}+b_{\mathbf{q}\lambda}\right)\left(b^\dagger{}_{-\mathbf{q}'\lambda'}+b_{\mathbf{q}'\lambda'}\right),
\end{aligned}
\tag{4.12}
$$

respectively. The derivatives of $\hat{V}_{\mathrm{scf}}$ with respect to the atomic coordinates **R** and phonons with energies $\{\omega_{\mathbf{q}\lambda}\}$ and polarization vectors $\{\xi_\alpha(\mathbf{q}\lambda|s)\}$ in Eqs (4.7) and (4.10) can be calculated by using density-functional perturbation theory (DFPT)[2].

Now we can start to take perturbations using Wick's theorem as the case of e-e interaction shown in subsection 2-2-3. For $\hat{H}_1$, pairing of the bosonic creation and

annihilation operators can be achieved only in the even-order perturbations. Thus the lowest order perturbation is the second. Since expectation values of the electronic and phonon parts are separately calculated, what we have to calculate is the following:

$$\begin{aligned}&\langle T\{c^{\dagger}_{n_1\mathbf{k}_1}(t_1)c_{n_1'\mathbf{k}_1-\mathbf{q}}(t_1)c^{\dagger}_{n_2\mathbf{k}_2}(t_2)c_{n_2'\mathbf{k}_2-\mathbf{q}'}(t_2)c_{n\mathbf{k}}(t)c^{\dagger}_{n\mathbf{k}}(t')\}\rangle \\ &\times\langle T\left\{\left(b^{\dagger}{}_{-\mathbf{q}\lambda}(t_1)+b_{\mathbf{q}\lambda}(t_1)\right)\left(b^{\dagger}{}_{-\mathbf{q}'\lambda'}(t_2)+b_{\mathbf{q}'\lambda'}(t_2)\right)\right\}\rangle\end{aligned} \tag{4.13}$$

T-product of phonon part in Eq (4.13) is written as follows:

$$\begin{aligned}&\langle T\left\{\left(b^{\dagger}{}_{-\mathbf{q}\lambda}(t_1)+b_{\mathbf{q}\lambda}(t_1)\right)\left(b^{\dagger}{}_{-\mathbf{q}'\lambda'}(t_2)+b_{\mathbf{q}'\lambda'}(t_2)\right)\right\}\rangle \\ &=\delta_{\mathbf{q},-\mathbf{q}'}\delta_{\lambda,\lambda'}\left[\langle T\{b^{\dagger}{}_{-\mathbf{q}\lambda}(t_1)b_{-\mathbf{q}\lambda}(t_2)\}\rangle+\langle T\{b_{\mathbf{q}\lambda}(t_1)b^{\dagger}{}_{\mathbf{q}\lambda}(t_2)\}\rangle\right] \\ &=\delta_{\mathbf{q},-\mathbf{q}'}\delta_{\lambda,\lambda'}\Big[\big[\theta(t_1-t_2)\mathrm{e}^{i\omega_{\mathbf{q}\lambda}(t_1-t_2)}N_{-\mathbf{q}\lambda}(T) \\ &\qquad+\theta(t_2-t_1)\mathrm{e}^{i\omega_{\mathbf{q}\lambda}(t_1-t_2)}\big(N_{-\mathbf{q}\lambda}(T)+1\big)\big] \\ &\qquad+\theta(t_1-t_2)\mathrm{e}^{-i\omega_{\mathbf{q}\lambda}(t_1-t_2)}\big(N_{\mathbf{q}\lambda}(T)+1\big) \\ &\qquad+\theta(t_2-t_1)\mathrm{e}^{-i\omega_{\mathbf{q}\lambda}(t_1-t_2)}N_{\mathbf{q}\lambda}(T)\Big] \\ &\equiv\delta_{\mathbf{q},-\mathbf{q}'}\delta_{\lambda,\lambda'}iD_0^{\mathbf{q}\lambda}(t_1-t_2,T).\end{aligned} \tag{4.14}$$

where $N_{\mathbf{q}\lambda}(T)$ is the Bose-Einstein distribution function of the phonon mode ($\mathbf{q}$, λ) at temperature *T*. $D_0^{\mathbf{q}\lambda}(t_1-t_2,T)$ is defined as free phonon Green's function. Now we focus on the case of zero-temperature:

$$iD_0^{\mathbf{q}\lambda}(t_1-t_2,T=0\text{ K})=\theta(t_2-t_1)\mathrm{e}^{i\omega_{\mathbf{q}\lambda}(t_1-t_2)}+\theta(t_1-t_2)\mathrm{e}^{-i\omega_{\mathbf{q}\lambda}(t_1-t_2)}.$$

T-product of the electronic part in Eq (4.13) is similar to that of the first-order perturbation of e-e interaction (Eq (2.61)). Self-energy corresponding to the Fock self-energy

represented in Eq (2.72) is called "Fan" self-energy:

$$\Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(t_1 - t_2, T = 0\ \mathrm{K}) = i \sum_{n'\mathbf{q}\lambda} \frac{\left|g_{nn'\mathbf{k}}^{\mathbf{q}\lambda}\right|^2}{N_q} \left[D_0^{\mathbf{q}\lambda}(t_1 - t_2, T = 0\ \mathrm{K}) G_0^{n'\mathbf{k}-\mathbf{q}}(t_1 - t_2)\right]. \quad (4.15)$$

After the Fourier-transform, this self-energy can be written in energy-space as follows,

$$\begin{aligned} &\Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(\omega, T = 0\ \mathrm{K}) \\ &= \sum_{n'\mathbf{q}\lambda} \frac{\left|g_{nn'\mathbf{k}}^{\mathbf{q}\lambda}\right|^2}{N_q} \left[\frac{1 - f_{n'\mathbf{k}-\mathbf{q}}}{\omega - \varepsilon_{n'\mathbf{k}-\mathbf{q}}^{\mathrm{KS}} - \omega_{\mathbf{q}\lambda} + i0^+} + \frac{f_{n'\mathbf{k}-\mathbf{q}}}{\omega - \varepsilon_{n'\mathbf{k}-\mathbf{q}}^{\mathrm{KS}} + \omega_{\mathbf{q}\lambda} - i0^+}\right], \end{aligned} \quad (4.16)$$

where $\varepsilon_{n'\mathbf{k}-\mathbf{q}}^{\mathrm{KS}}$ is Kohn-Sham energy of the *n'*th band at the point **k**−**q** in the Brillouin zone. $f_{n'\mathbf{k}-\mathbf{q}}$ is the occupation number of the bare electronic state at (n', **k**−**q**). Using Matsubara Green's function,[1] the Fan self-energy can be defined in finite-temperature case:

$$\begin{aligned} &\Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(\omega, T) \\ &= \sum_{n'\mathbf{q}\lambda} \frac{\left|g_{nn'\mathbf{k}}^{\mathbf{q}\lambda}\right|^2}{N_q} \left[\frac{N_{\mathbf{q}\lambda}(T) + 1 - f_{n'\mathbf{k}-\mathbf{q}}}{\omega - \varepsilon_{n'\mathbf{k}-\mathbf{q}}^{\mathrm{KS}} - \omega_{\mathbf{q}\lambda} - i0^+} + \frac{N_{\mathbf{q}\lambda}(T) + f_{n'\mathbf{k}-\mathbf{q}}}{\omega - \varepsilon_{n'\mathbf{k}-\mathbf{q}}^{\mathrm{KS}} + \omega_{\mathbf{q}\lambda} - i0^+}\right]. \end{aligned} \quad (4.17)$$

Similarly, the Debye-Waller (DW) self-energy arises from the first-order term in the perturbative expansion in powers of $\hat{H}_2$,

$$\Sigma_{n\mathbf{k}}^{\mathrm{DW}}(T) = \frac{1}{N_q} \sum_{\mathbf{q}\lambda} \Lambda_{nn\mathbf{k}}^{\mathbf{q}\lambda,-\mathbf{q}\lambda} \left(2N_{\mathbf{q}\lambda}(T) + 1\right). \quad (4.18)$$

Explicitly both self-energies depend on the temperature *T* via the $N_{\mathbf{q}\lambda}(T)$ factor. The Fan and DW self-energies are complex and real functions, respectively; thus the former gives both an e-ph-induced energy shift and broadening while the latter contributes only with a constant energy shift.

Using two self-energies, the fully interacting Green's function $G_{nk}(\omega, T)$ can be written as

$$G_{n\mathbf{k}}(\omega, T) = \frac{1}{\omega - \varepsilon_{n\mathbf{k}}^{\mathrm{KS}} - \Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(\omega, T) - \Sigma_{n\mathbf{k}}^{\mathrm{DW}}(T)}. \quad (4.19)$$

In the QPA described in subsection 2-2-5, the energy of quasiparticle, $E_{n\mathbf{k}}(T)$, is written as

$$E_{n\mathbf{k}}(T) = \varepsilon_{n\mathbf{k}}^{\mathrm{KS}} + Z_{n\mathbf{k}}(T)\left[\Sigma_{n\mathbf{k}}^{\mathrm{Fan}}\left(\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}, T\right) + \Sigma_{n\mathbf{k}}^{\mathrm{DW}}(T)\right]. \quad (4.20)$$

The on-mass shell (OMS) approximation represents the static limit of the QPA, obtained by assuming $\Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(\omega, T) \approx \Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(\omega, T)\big|_{\omega=\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}}$, which is equivalent to assume $Z_{n\mathbf{k}}(T) = 1$ in Eq (4.20).

The key importance of considering dynamical effect has been recently discovered.[3] For instance, diamond has been shown to have large dynamical EP effects, which explains the subgap states observed experimentally in the absorption spectrum. Similarly, carbon polymer systems like *trans*-polyacetylene and polyethylene, show a severe breakdown of the QP picture induced by the EP interaction. In such cases, Eq (4.20) based on the QPA is not correct.

4-2. Finite-temperature electron-phonon correlation effects on the electronic and optical properties of zinc-blende GaN

4-2-1. Introduction to zinc-blende GaN

The group III-nitride semiconductors, i.e., GaN, AlN, InN, and their alloys are materials with many applications in the field of optoelectronics. These include, among others, light emitting diodes (LEDs), laser diodes (LDs), and heterojuction field-effect transistors (HFETs).[4–8] This class of compounds is widely used, being characterized by the most stable wurtzite structure. They have built-in electric fields arising from the spontaneous and piezoelectric polarization along the *c* axis. These fields are, however, undesirable in the application of the heterostructures as quantum wells (QWs) or superlattices since they complicate the design and worsen the sample malleability. One of the approaches to eliminate these internal fields is the use of metastable nonpolar zinc-blende (*zb*) structures. It has also been reported that *zb* group III nitrides have a quantum confined Stark effect in low-dimensional heterostructures,[9] high *p*-type conductivity in (Ga,Mn)N thin films,[10] and negative differential resistance (NDR) at the resonant tunneling diode of the cubic Al(Ga)N/GaN.[11,12] Consequently a lot of interest is constantly attracted by this family of materials.

In the last few years *zb*-GaN with high phase-purity and crystalline quality has been fabricated as a nearly strain-free epitaxial layer on 3C-SiC(001)/Si pseudosubstrates by

plasma assisted molecular beam epitaxy.[9,12–14] This experimental achievement boosted the interest on fundamental optical properties such as photoluminescence, photoreflectance, and ellipsometry with particular attention on their temperature dependence. In contrast to such abundance of experimental results the agreement with the state-of-the-art calculations of the optical properties of *zb*-GaN is still not satisfactory. In these approaches the absorption spectrum is calculated[15] by including electron-hole interaction by solving the Bethe-Salpeter equation (BSE) derived within the MBPT.[16] Nevertheless, the main peak position is strongly underestimated when compared to the experimental result. Also the complex temperature dependence observed experimentally is not captured at all. Similarly, the band structure of *zb*-GaN has been deeply investigated by using the most up-to-date theoretical approaches. In this case electron-electron correlation only has been included, by means of the well-known GW approximation.[17] The corresponding quasiparticle (QP) gap, calculated by using the one-shot GW approximation on top of Kohn-Sham (KS) Heyd-Scuseria-Ernzerhof (HSE)[18] hybrid orbital (HSE+G_0W_0), is 3.427 eV,[19] which overestimates the experimental value of 3.295 eV.[9]

In this work I calculate the electronic and optical properties of *zb*-GaN by including the e-ph and e-e interactions. My results show a remarkable impact of e-ph interaction even at zero temperature which corrects the overestimation of the QP gap obtained within the HSE+G_0W_0 method. At the same time I prove that only by treating on the same level e-e and

e-ph interactions is it possible to obtain an unprecedented agreement with experiment results, both at zero temperature and at finite temperature.

4-2-2. Renormalization of single-particle energy levels

zb-GaN is a polar material and, as a consequence, large static e-ph effects are expected.[20] As mentioned above, a strong e-ph coupling might eventually induce the breakdown of the QPA. A clear and simple way to test the QPA validity is to calculate the renormalization factors $Z_{n\mathbf{k}}(T)$. Indeed, within the QPA, the Green's function G can be written as

$$G_{nk}^{\mathrm{QP}}(\omega, T) = \frac{Z_{n\mathbf{k}}(T)}{\omega - E_{n\mathbf{k}}(T)}, \tag{4.21}$$

with $E_{n\mathbf{k}}(T)$ evaluated by means of Eq. (4.20). The imaginary part of the Green's function $A_{n\mathbf{k}}(\omega, T) \equiv \pi^{-1}|\mathrm{Im}[G_{n\mathbf{k}}(\omega, T)]|$ gives the electronic spectral function (SF). In the OMS approximation ($Z_{n\mathbf{k}}$ = 1), the SF reduces to a Lorentzian function with a pole at $\omega = \mathrm{Re}[E_{n\mathbf{k}}(T)]$ and width $\Gamma_{n\mathbf{k}}(T) = \mathrm{Im}[E_{n\mathbf{k}}(T)]$. Thus, the $Z_{n\mathbf{k}}$ values measure the strength of the QP pole; i.e., the QP picture is well motivated when the SF can be well approximated with a single Lorentzian-like function.

In my e-ph calculations the optimized geometry and the electronic state are obtained by using the PWscf code.[21] EP calculations are performed with the YAMBO code[22] by using the phonons frequencies and $g_{nn'\mathbf{k}}^{\mathbf{q}\lambda}$ matrix elements calculated with PWscf within DFPT. An energy cutoff on the plane-wave expansion of 80 Ry and a uniform k-point sampling of 8 × 8

× 8 are used in the ground-state and DFPT calculations. In order to obtain converged e-ph self-energies, a large number of **q**-points and unoccupied bands are required. Therefore I used 400 bands to evaluate the energy shift (related to $\mathrm{Re}[E_{n\mathbf{k}}(T)]$) and 700 randomly generated **q**-points for the broadening (linked to the $\mathrm{Im}[E_{n\mathbf{k}}(T)]$), respectively. As a result of my simulations the majority of the states that contribute to the optical absorption are well described by Lorentzian-like SF, as shown in Fig. 4-1. In addition most of the states show values of $Z_{n\mathbf{k}}$ very close to 1. For example, the states corresponding to the valence band maximum and the conduction band minimum at the point have $Z_{n\mathbf{k}} = 0.91$ and $Z_{n\mathbf{k}} = 0.98$, respectively.

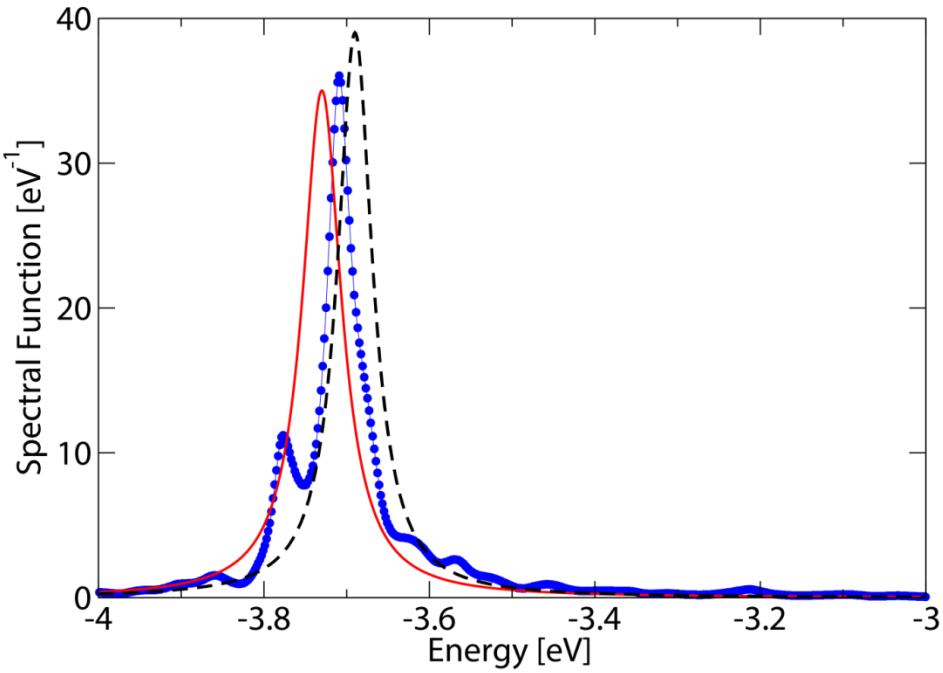


Figure 4-1. Spectral function of a valence band state. This well represents the general trend of the majority of spectral functions covering the energy range involved in the absorption process, as discussed in the text. The blue line with dots is the calculated SF. This is compared with two Lorentzian functions corresponding to the OMS (red solid line) and to the QPA (black dashed line). Both approximations reproduce the calculated SF quite well and the use of the OMS is, therefore, well motivated.

This indicates that, in *zb*-GaN, the OMS approximation is well motivated and most of the weight can be safely assumed to be in one single peak. There is, however, another and more stringent motivation in favor of OMS as far as the calculation of the optical properties is concerned. A $Z_{n\mathbf{k}}$ factor smaller than 1 is known to reduce the intensity of the absorption spectrum. At the same time, however, it is well known that such reduction is compensated by

the dynamical electron-hole interactions.[23,24] As far as these dynamical effects are neglected (as is commonly done in the state-of-the-art implementation of the BSE used in this work) the OMS assumption of $Z_{n\mathbf{k}} = 1$ is well motivated also from a purely theoretical point of view.

In order to describe the impact of the e-ph interaction on the electronic states I consider the energies corresponding to the lowest transitions at several high-symmetry points. These energies are compared with the experimental results in Table 4-1. DFT is well known to underestimate the band gaps by about 40%. In fact, my DFT calculation [performed with the local density approximation (LDA)] yields 2.231 eV as the band gap of *zb*-GaN. This is clearly less than the experimental value, that is 3.295 eV at 10 K.[9] My LDA+G_0W_0 calculation within the plasmon-pole approximation[17] opens the gap to 3.239 eV, which well agrees with the experiment. Still, however, the LDA+G_0W_0 largely underestimates the transition energies at L and X. In my G_0W_0 calculation the energy cutoff in the Fourier expansion of the response function is 13 Ry and 80 bands are included. This underestimation can be traced back to the local treatment of e-e correlation effects in the self-consistent DFT-LDA calculation. This limitation can be overcome by using the AM05 functional[25] to calculate the optimized geometry and the HSE functional[18] for the start point of the G_0W_0 calculation, as previously reported.[19]

Table 4-1. Lowest transition energies at high-symmetry points in the Brillouin zone of *zb*-GaN. The values obtained from LDA, LDA+G_0W_0, HSE+ G_0W_0,[19] and HSE+G_0W_0+OMS calculations are compared with the experimental values.[9] All values are in eV.

	Γ	L	X
LDA	2.231	5.952	6.034
LDA+G_0W_0	3.239	7.117	7.105
HSE+G_0W_0	3.427	7.707	7.755
HSE+G_0W_0+OMS	3.300	7.517	7.624
Exp (T = 10K)	3.295	7.33	7.62

[a] Excitation peak positions including the electron-hole binding energies.

From Table 4-1 it is evident that HSE+G_0W_0 opens the band gaps further and overestimates the transition energies at all three high-symmetry points[19]. However, the combination of e-ph interaction with HSE+G_0W_0 greatly compensates this overestimation leading to an excellent agreement with the experimental results. My calculation at the OMS level gives a gap correction of −0.127 eV at, which reduces the HSE+G_0W_0 gap to 3.300 eV, in agreement with the experiment. Similarly, at the L and X points, the EP-induced corrections are −0.190 and −0.132 eV, resulting in transition energies of 7.517 and 7.624 eV, again in very good agreement with the experiment. Such moderate e-ph corrections result from the large cancellation between the Fan and the DW contributions to the total self-energy. For instance, the gap correction −0.127 eV at the point is decomposed into a Fan contribution (+2.101 eV)

and a DW contribution (−2.228 eV). This large cancellation clearly points to the importance of including, at the same time, both contributions to the total self-energy. Moreover, my result indicates the significance of the EP correction in *zb*-GaN, pointing to similar and potentially important corrections in the whole III-nitrides group of materials.

4-2-3. Finite-temperature optical absorption spectra

The optical absorption spectrum is defined as the imaginary part of the macroscopic dielectric function $\mathrm{Im}[\varepsilon_{\mathrm{M}}(\omega)]$. This can be easily expressed, in the long wavelength limit, as

$$\varepsilon_{\mathrm{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0} v_0(\mathbf{q}) \int d\mathbf{r} d\mathbf{r}' \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{r}-\mathbf{r}')} \bar{P}(\mathbf{r},\mathbf{r}';\omega) \qquad (4.22)$$

with $v_{\mathbf{G}}(\mathbf{q}) = \frac{4\pi}{|\mathbf{G}+\mathbf{q}|^2}$ being the Coulomb potential and $\bar{P}(\mathbf{r},\mathbf{r}';\omega)$ the two-point polarizability. The equation of motion for the polarizability[16] can be rewritten by introducing a single-particle basis set ($\{\psi_{n,\mathbf{k}}\}$) to expand the density operator. This is equivalent to define the electron-hole probability functions $\Phi_{\mathbf{K}}(\mathbf{r}) = \psi_{c\mathbf{k}}(\mathbf{r})\psi^*_{v\mathbf{k}}(\mathbf{r})$. Here $\mathbf{K}$ represents the general conduction-valence pairs, $\mathbf{K} = (c, v, \mathbf{k})$. In this basis $\bar{P}$ is

$$\bar{P}(\mathbf{r},\mathbf{r}';\omega) = -\left(\frac{i}{\Omega N}\right) \sum_{\mathbf{K}_1\mathbf{K}_2} \Phi^*_{\mathbf{K}_1}(\mathbf{r}) L_{\mathbf{K}_1\mathbf{K}_2}(\omega) \Phi_{\mathbf{K}_2}(\mathbf{r}') \qquad (4.23)$$

Eq (4.23) introduces the electron-hole Green's function $L_{\mathbf{K}_1\mathbf{K}_2}(\omega)$ that satisfies the BSE equation[16]:

$$L_{\mathbf{K}_1\mathbf{K}_2}(\omega) = L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega) + L^0_{\mathbf{K}_1\mathbf{K}_3}(\omega) \Xi_{\mathbf{K}_3\mathbf{K}_4}(\omega) L_{\mathbf{K}_4\mathbf{K}_2}(\omega) \qquad (4.24)$$

The Bethe-Salpeter kernel Ξ is defined as $\Xi = -iV + iW$, with V and W being the exchange

and the screened Coulomb interactions, respectively. $L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega)$, in Eq. (4.24), is the free electron-hole Green's function.

As previously described by Marini[26] it is possible to include the finite-temperature effect in the BSE by using as reference the single-particle energies and the temperature-dependent and complex QP energies $E_{n\mathbf{k}}(T)$. In this way the free electron-hole Green's function $L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega, T)$ depends explicitly on the temperature:

$$L^0_{\mathbf{K}_1\mathbf{K}_2}(\omega, T) = i\left[\frac{f_{c_1\mathbf{k}_1} - f_{v_1\mathbf{k}_1}}{\omega - E_{c_1\mathbf{k}_1}(T) + E_{v_1\mathbf{k}_1}(T) + i0^+}\right]\delta_{\mathbf{K}_1\mathbf{K}_2} \qquad (4.25)$$

Eq (4.25) ensures that also the fully interacting electron-hole Green's function and the absorption spectra depend explicitly on the temperature, thanks to Eqs (4.22) and (4.23).

In order to solve the BSE I adopt two standard approximations. The first is the Tamm-Dancoff approximation which corresponds to neglecting the coupling between the resonant and the antiresonant part of the BSE kernel. The second is the use of the statically screened electron-hole potential *W*.

In Fig. 4-2 I show the calculated absorption spectrum. In addition to the G_0W_0 corrections, I include e-ph effects at the OMS level. To obtain converged absorption spectra I employed the random-integration method (RIM)[22] by selecting around 30,000 random **k** points in the whole Brillouin zone. The energy cutoffs for the exchange and screened Coulomb interactions are set to 60 and 3 Ry, respectively. The artificial damping parameter $0+$ in Eq. (4.25), which is introduced only for numerical reasons, is 10 meV. The resulting spectrum (thick black line)

is compared to the previous calculation of Benedict and Shirley[15] (dashed red line) which is performed in a LDA basis without including the e-ph interaction. The BSE spectrum, obtained by including only the HSE+G_0W_0 correction, is also shown (blue line).

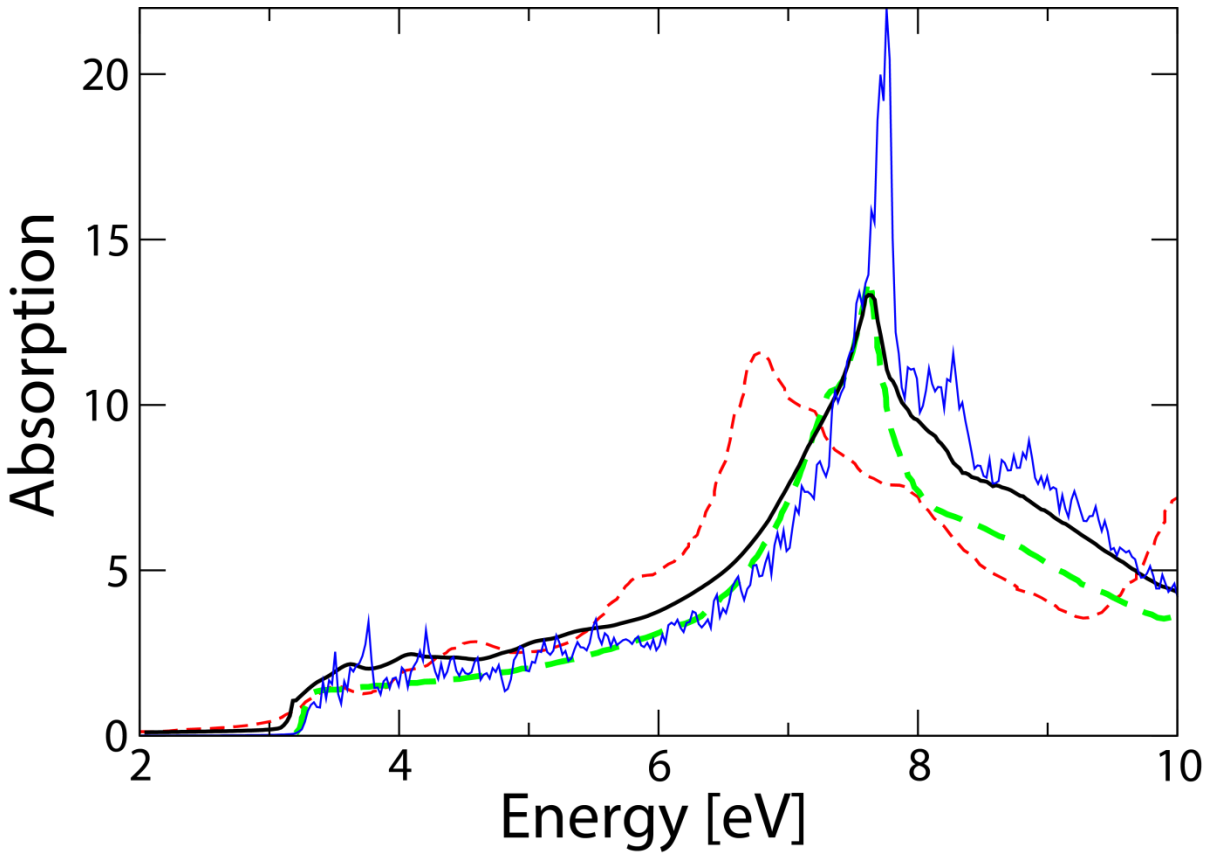


Figure. 4-2. Theoretical and experimental absorption spectra of *zb*-GaN at T = 10 K. The spectrum obtained by solving the BSE including both HSE+G_0W_0 and e-ph corrections (thick black line) is in excellent agreement with the experimental results[9] (bold dashed green line). Compared to the state-of-the-art calculation of Benedict and Shirley[15] (dashed red line), the agreement is largely improved. A blue line shows the BSE spectrum only with the HSE+G_0W_0 correction.

I notice, indeed, that, compared to the Benedict's calculation, the calculated position and width of the main peak are in very good agreement with the experimental spectrum at T = 10 K[9] (bold dashed green line). This pronounced peak, located at around 7.62 eV in the experiment, is due to interband transitions in a region of the Brillouin zone near the X point. These transitions extend over regions where the valence and conduction bands are parallel with a similar energy distance.[9] Therefore the present improvement is largely due to the correction from HSE+G_0W_0 that induces a large stretching of the bands. In the BSE spectrum only with HSE+G_0W_0, the peak position is clearly shifted to higher energies and it is much less energetically wide.

Indeed, one of the most significant effects of the EP interaction is the broadening of the QP states that fully dictates the smooth energy dependence of the absorption spectrum. I use a very small artificial damping (10 meV) in Eq. (4.25) that leads, when the EP interaction is neglected, to a very spiky absorption spectrum. To avoid this unphysical behavior an arbitrary 200 meV broadening is chosen in the previous BSE calculations. In the present case, instead, the effect of the e-ph interaction correctly describes both the main peak and the steep absorption edge. Finally I investigate how the optical spectrum evolves as the temperature is increased.

In Fig. 4-3 I show the calculated absorption spectra at T = 0, 300, and 600 K. In the first two cases the numerical simulation is compared with the available experimental results.[9] The

agreement is fairly good and confirms that the present approach is able to correctly capture the finite-temperature effects. The energy shifts of the first excitation peaks due to electron-hole transitions occurring at Γ, L, and X are experimentally −63, −100, and −110 meV, respectively, when the temperature is increased from 10 to 295 K.[9] My calculations give −83, −144, and −106 meV, respectively, which show a quantitative agreement with experiment. Also the general trend observed experimentally that the peak shift is larger at critical points with higher transition energy is reproduced.

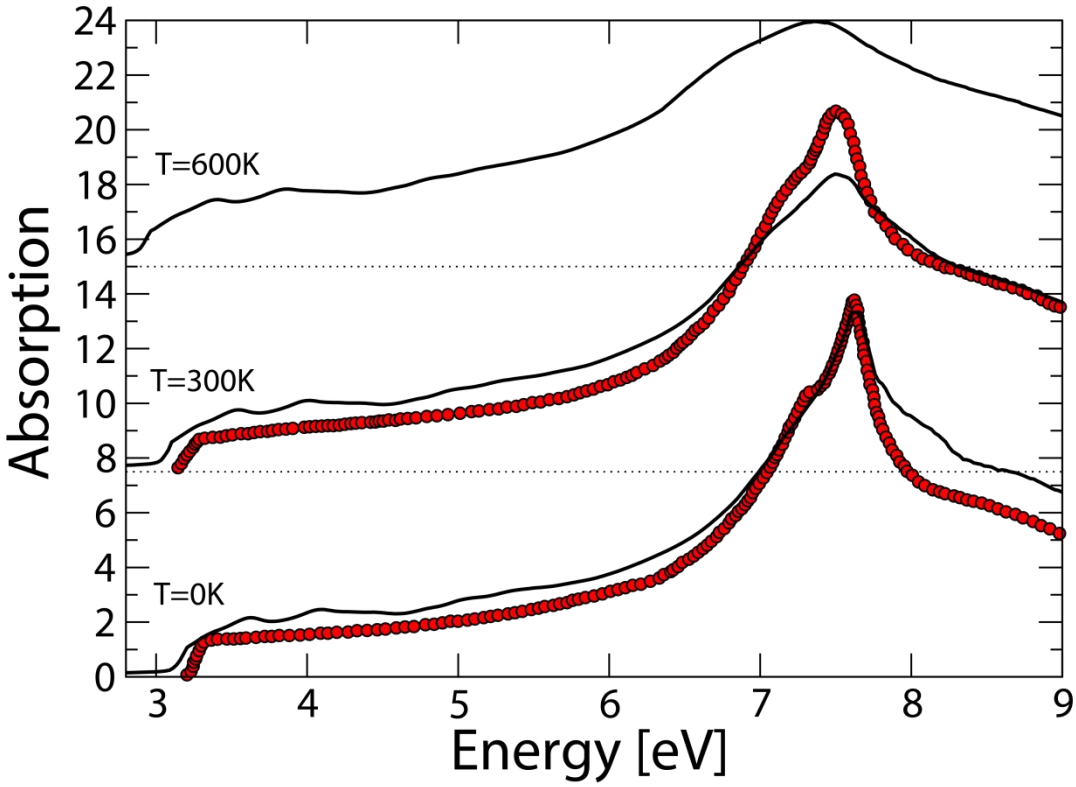


Figure 4-3. Absorption spectra of *zb*-GaN at T = 0, 300, and 600 K. Red circles are experimental results at 10 and 295 K.[9]

On the other hand, from Fig. 4-3 I deduce that the broadening of the main peak at the X point is slightly overestimated at T = 300 K compared to the experiment. In order to understand the source of this overestimation I notice that, in the QP picture, the e-ph-induced broadening of the valence band top and the conduction band bottom at the X point is 121.8 and 12.1 meV, respectively, at T = 300 K. In the independent particle approximation (where $L_{\mathbf{K}_1\mathbf{K}_2} \approx L^0_{\mathbf{K}_1\mathbf{K}_2}\delta_{\mathbf{K}_1\mathbf{K}_2}$) the electron-hole broadening is simply the sum of the two. Now, as also in the case where electron-hole attraction is included, the main absorption peak originates from the lowest transitions concentrated around X. Thus I deduce that the overestimation is due to too large broadening of the underlying QP states. In the experimental work by Logothetidis *et al.*[27] the broadening at the main absorption peak is described by a phenomenological model,

$$\Gamma(T) = \Gamma_1 + \Gamma_0\left[1 + \frac{2}{\exp(\Theta/T) - 1}\right] \qquad (4.26)$$

with Γ_1 = 27 meV, Γ_0 = 44 meV, and Θ = 522 K. The first term in Γ_1 is ascribed to a temperature-independent mechanism, such as surface scattering; thus I set it to 0 to be compared with my theoretical results. Equation (4.26), indeed, predicts $\Gamma(T = 300\ \text{K}) \sim 62.7\ \text{meV}$ that is half of the value of my results.

A reasonable explanation of this deviation is in the underlying unperturbed band structure. The band curvature has a large impact on the EP-induced broadening through the denominator of Eq. (4.17), especially by the dominant intraband-scattering terms with

$\omega = \varepsilon_{n\mathbf{k}}^{\mathrm{KS}}$, $n' = n$, and small **q**. Since my e-ph self-energies are calculated on top of Kohn-Sham states from the LDA, the resulted band widths are too small. As shown in previous calculations[19] the valence band at the X point is characterized by a large curvature that is underestimated by the LDA calculations. I expect that the broadening would be improved by e-ph calculations performed on top of the HSE+G_0W_0 band structure, but it is prohibitively expensive from the computational point of view. Nevertheless my approach, based on the LDA, gives excellent results especially at the low temperature.

4-2-4 Conclusion

In this work I study the zero- and finite-temperature electronic and optical properties of *zb*-GaN. The effect of e-ph interaction, treated in a fully dynamical approach based on the MBPT, shows that the simple on-the-mass-shell approximation to the QP energies and widths is well motivated for the low-energy states involved in the absorption spectrum.

By including, in an *ab initio* manner, the combined effect of the e-e interaction and the e-ph interaction I obtain an excellent agreement with the experimental fundamental band gaps. The solution of the BSE calculated on top of the HSE + G_0W_0 band structure including e-ph effects leads to an excellent agreement also for the optical absorption spectrum measured on high phase-purity samples. Both the position and the broadening of the most intense absorption peak are correctly reproduced in the low-temperature regime. In the

room-temperature case, instead, the red-shift of the main peak position is well described while the broadening is slightly overestimated. Despite this overestimation the present results still represent a major improvement with respect to the state-of-the-art simulations.

My results clearly point to the crucial importance of including at the same time electron-electron and electron-phonon correlation effects for a comprehensive and quantitative understanding of the electronic and optical properties of group III nitrides.

4-3. The mechanism of slow hot-hole cooling in lead-iodide perovskite

4-3-1. Introduction to photovoltaic material organic-inorganic halide perovskite

Over the last five years, a new class of photovoltaic (PV) cells based on organic–inorganic mixed halide perovskites (ABX_3, where A = alkali/organic cation, B = Sn, Ge, Pb, and X = halide anion) has emerged as solar harvesters. Moving on from the first report by Kojima *et al.* in 2009, where liquid electrolyte-based dye-sensitized solar cells (DSSCs) employing $CH_3NH_3PbI_3$ and $CH_3NH_3PbBr_3$ were characterized as having a power conversion efficiency (PCE) of about 3%,[28] perovskite solar cells have improved rapidly and nowadays the PCE exceeds 15% in solid state solar cell architectures.[29–31]

Achieving both good photoabsorption properties and a high carrier mobility is a difficult task in solution-processed PVs, organic PVs, DSSCs, and colloidal quantum dot solar cells because of the poor crystallinity of these absorbers.[32] One of the main advantages of organic–inorganic halide perovskite materials stems from the high mobility of both their electrons and their holes (i.e., their ambipolar character[33–40]), which can be advantageously coupled with their low-cost manufacturing for PV applications. Recent studies on their photoluminescence (PL) and time-resolved transient absorption spectra (TAS) have focused on the carrier

lifetimes and diffusion lengths of $CH_3NH_3PbI_3$.[35,36] TAS measurements by Xing *et al*. on $CH_3NH_3PbI_3$ have revealed that the diffusion lengths of the electrons and holes are 130 and 110 nm, respectively.[35]

A long hot-hole lifetime in $CH_3NH_3PbI_3$ was reported as a secondary product of the TAS analysis. The UV–Vis absorbance spectrum of $CH_3NH_3PbI_3$ shows two absorption peaks occurring at 760 nm and 480 nm. The former peak was attributed to the direct excitation from the valence band maximum (VBM), which Xing *et al*. denoted as VB1, to the conduction band minimum (CBM). The latter peak was attributed to the transition from a band (VB2) situated 0.9 eV below VB1 to the CBM.[35] Xing *et al*. reported that the hot-hole cooling from VB2 to VB1 was a very slow process (with a lifetime of ~0.4 ps). A suggested scheme of the experimentally observed carrier dynamics is shown in Fig. 4-4. The lifetime of this process is much longer than that typically seen in most organic semiconductors (~100 fs). As Xing *et al*. proposed, these hot-hole energies can be efficiently extracted through optimizing the device configuration, and may pave the way to further improvements in conversion efficiency.

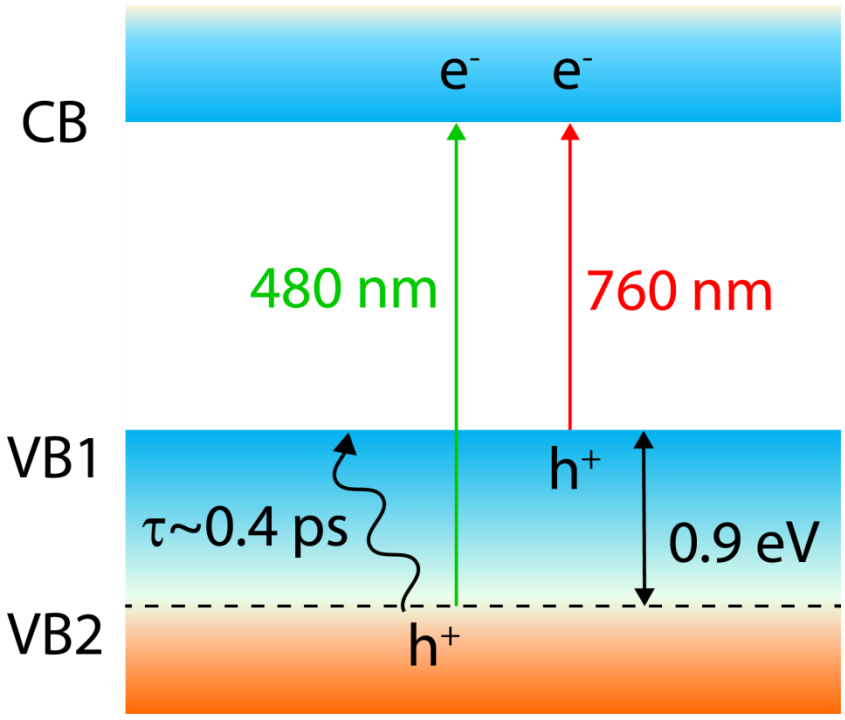


Figure 4-4. A schematic diagram showing hot-hole cooling in $CH_3NH_3PbI_3$ as reported in Ref. [35].

This study aims to reveal the origin of slow hot-hole cooling using first-principles calculations from the comparison of two lead iodide perovskite materials $CH_3NH_3PbI_3$, $CsPbI_3$, and the charged semiconductor network, PbI_3^-. In semiconductors, the most relevant mechanism of intraband carrier relaxation is thermalization coupled with lattice vibrations, i.e., e-ph interactions. As shown in the previous sections, this is what we can calculate in MBPT scheme. Using *ab initio* calculations, I have investigated the e-ph lifetimes $\tau_{n\mathbf{k}}$ defined as

$$\tau_{n\mathbf{k}}^{-1} \equiv 2\mathrm{Im}E_{n\mathbf{k}} \approx 2\mathrm{Im}\Sigma_{n\mathbf{k}}^{\mathrm{Fan}}(\varepsilon_{n\mathbf{k}}^{\mathrm{KS}}) \tag{4.27}$$

for lead iodide perovskite materials.

4-3-2. Band structure and phonon band

My initial structural optimizations were performed using a DFT approach, as implemented in the PWscf code in the Quantum ESPRESSO software package.[21] The local density approximation (LDA) with the parameterization given by Perdew and Zunger[41] was used for the exchange correlation functional. Norm-conserving pseudopotentials, along with the plane wave basis set, were used. The cutoff energies for $CH_3NH_3PbI_3$, $CsPbI_3$, and PbI_3^- were set to 80, 60, and 40 Ry, respectively. In all the systems, the Brillouin zone was sampled using Γ-centered $6 \times 6 \times 6$ *k* point meshes employing the Monkhorst–Pack scheme.[42]

The optimized structures of the $CH_3NH_3PbI_3$, $CsPbI_3$, and PbI_3^- systems are shown in Fig. 4-5. Both $CsPbI_3$ and PbI_3^- retained their cubic structure, while the geometry of $CH_3NH_3PbI_3$ was distorted because of the anisotropic character of the $CH_3NH_3^+$ cation. The lattice constants and band gaps of these compounds are summarized in Table 4-2.

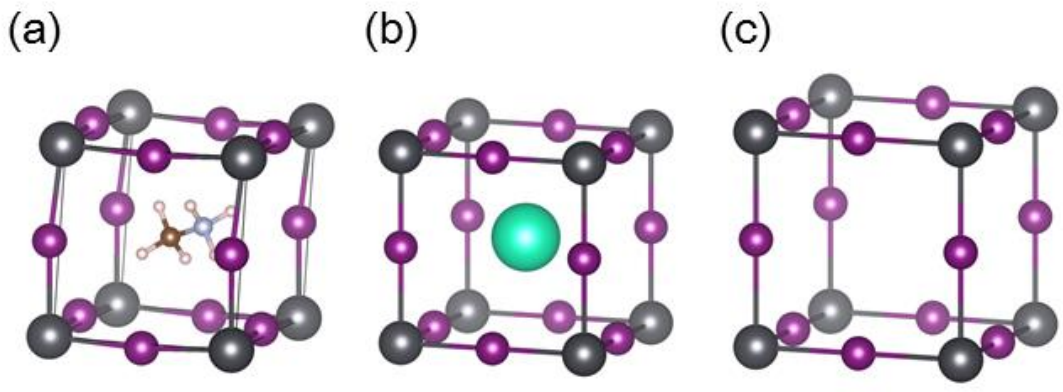


Figure 4-5. Optimized structures of (a) $CH_3NH_3PbI_3$, (b) $CsPbI_3$, and (c) PbI_3^-. Key: black = Pb, purple = I, pink = H, brown = C, light blue = N, and green = Cs.

Table 4-2. The lattice constants and band gaps of lead iodide perovskites optimized using LDA calculations.

	$CH_3NH_3PbI_3$	$CsPbI_3$	PbI_3^-
Lattice constant /Å	a=5.989, b=6.018, c=5.986 (α=86.52, β=90.96, γ=86.54)	6.133	6.721
Band gap /eV	0.926	0.894	1.509

The lattice constants of $CH_3NH_3PbI_3$ and $CsPbI_3$ were similar, and were smaller than those of PbI_3^-. Consistent with our previous results,[38] the absence of any Coulomb interactions resulting from the removal of the A-site cation resulted in an expansion of the cell volume.

The expansion of the cell size of PbI_3^- markedly altered the band structure. Fig. 4-6 shows the band structure and density of states (DOS) of the three systems. The band structures of $CH_3NH_3PbI_3$ and $CsPbI_3$ are almost identical, although a distortion in $CH_3NH_3PbI_3$ resolves the degeneracy. This is because the valence band and conduction band are mainly derived from the I 5p, Pb 6s, and Pb 6p orbitals in these lead halide perovskite systems.[37,38] Thus, different A-site cations only marginally impact on the electronic structure.

Fig. 4-6(a) and (b) show large band dispersions for both the VBM and the CBM, resulting in light effective masses of both electrons and holes in the CBM and VBM, which agrees with the experimentally reported ambipolar nature of $CH_3NH_3PbI_3$. However, PbI_3^- differs markedly from the two cases discussed above, as the band dispersion in the VBM is

manifestly reduced, inducing a heavier hole effective mass. This is because the antibonding coupling between the Pb 6s and I 5p orbitals is reduced in the expanded cell.[38,43] As a result, the DOS of $CH_3NH_3PbI_3$ and $CsPbI_3$ is smakk in the -1 eV $< E-E_{VBM} <$ 0 eV region (encircled area), while the DOS of PbI_3^- shows a noticeably high peak. I will show below how this difference has a marked impact on the e-ph lifetimes.

Since the computational burden of the phonon and e-ph calculations of $CH_3NH_3PbI_3$ is sensitively higher than that for $CsPbI_3$ and PbI_3^- according to the larger size ($CH_3NH_3PbI_3$, $CsPbI_3$, and PbI_3^- have 12, 5, and 4 atoms in the unit cell.), in this work I report the e-ph lifetimes of $CsPbI_3$ and PbI_3^-. Through a comparison of the e-ph lifetimes of $CsPbI_3$ and PbI_3^-, I was able to demonstrate that the origin of the slow hot-hole cooling was almost entirely from the characteristic electronic structure. Thus, my results for $CsPbI_3$ could be easily transferred to $CH_3NH_3PbI_3$.

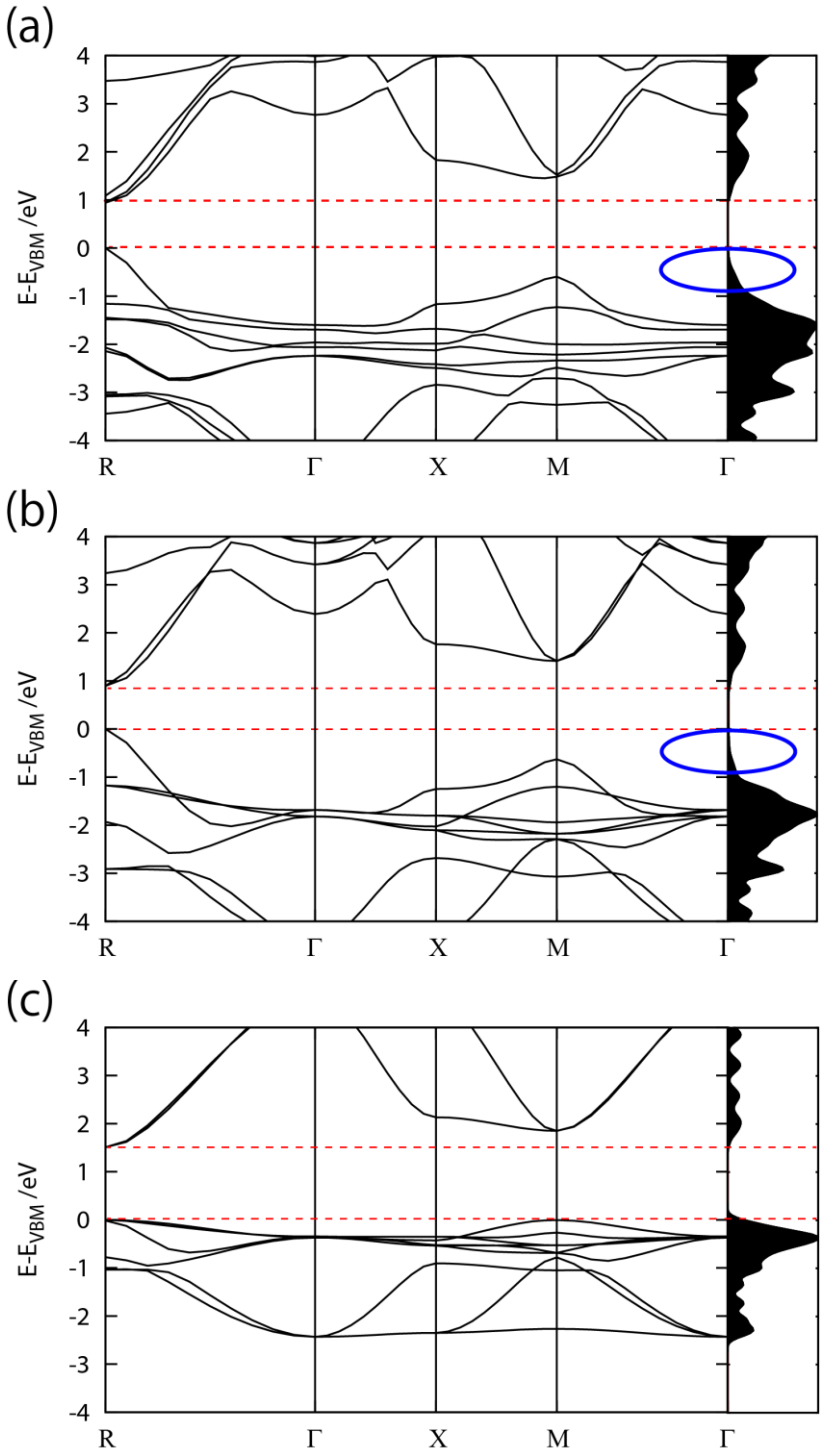


Figure 4-6. Bandplots (left) and DOS (right) of (a) $CH_3NH_3PbI_3$, (b) $CsPbI_3$, and (c) PbI_3^-.

The red dotted lines show the VBM and CBM.

The phonons of $CsPbI_3$ and PbI_3^- were calculated using the DFPT scheme (Fig. 4-7). I checked the phonon bands and found that the Debye frequencies of these two systems are 138 and 77 cm^{-1} for $CsPbI_3$ and PbI_3^-, respectively, corresponding to very small energies compared with the magnitude of the electronic energy scale (~1 eV). The number of real phonon bands was 9 and 10 for $CsPbI_3$ and PbI_3^-, respectively, while the number of all phonon modes was 15 and 12, respectively. The imaginary modes arise from the fact that the cubic structure is not the most stable phase in these systems: the most stable phase of $CsPbI_3$ is the δ-phase exhibiting an orthorhombic crystal system.[44] However, the number of atoms in the unit cell of the δ-phase is 20, which is four times that of the cubic phase. To the best of my knowledge, there have been no reports on *ab initio* calculations on the e-ph lifetimes in such a large system due to the computational overhead. Thus, at first, I here focus on the cubic polymorph discussing in the following the effect of imaginary phonon modes on the e-ph lifetimes.

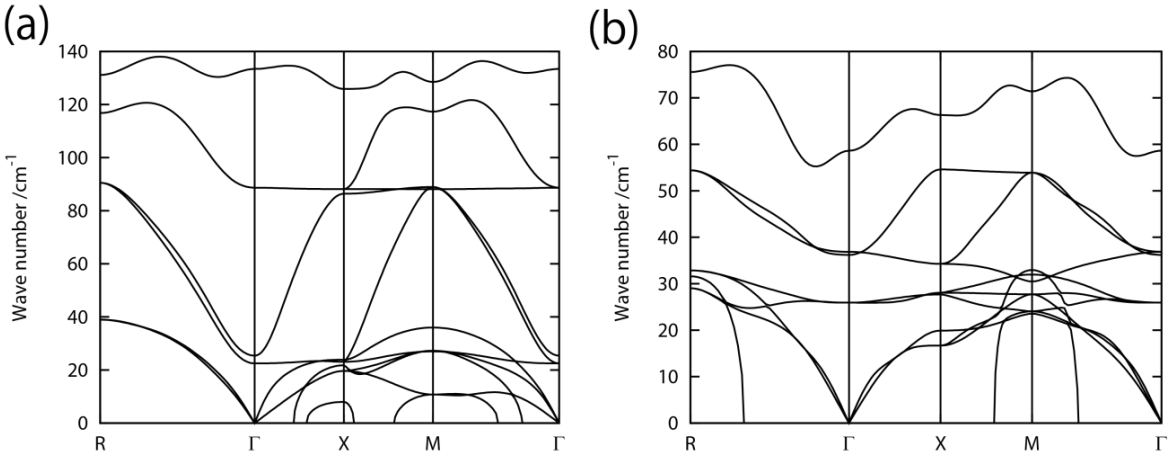


Figure 4-7. Phonon bands of (a) $CsPbI_3$ and (b) PbI_3^-

4-3-3. Carrier lifetimes

The resulting e-ph lifetimes of $CsPbI_3$ and PbI_3^- are plotted in Fig. 4-8, along with the KS energies. I also used the Phonon software package to calculate the e-ph coupling matrix. Then, MBPT calculations were performed to obtain the self-energies and the lifetimes employing the YAMBO code.[22] Here, the *k*-points were sampled in 8 × 8 × 8 grids and 30 bands were included in the self-energy calculations. The temperature was set to T = 0 K for simplicity.

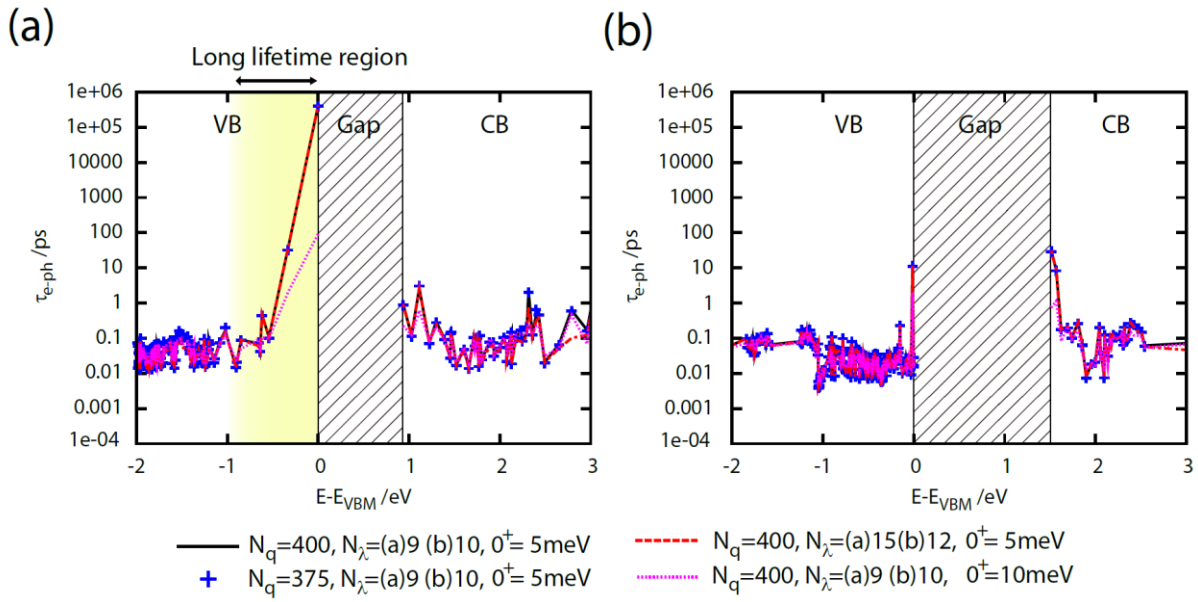


Figure 4-8. The e-ph lifetimes of (a) $CsPbI_3$ and (b) PbI_3^-. The black solid, red dashed, blue dashed, and pink dotted lines are from different calculation settings, as shown at the bottom of the figure.

The number of **q** points, $N_{\mathbf{q}}$ and damping parameter, 0^+ were crucial parameters that needed to be checked carefully in the convergence test procedure. The black solid lines in Fig.

4-8 correspond to the set of data with $N_{\mathbf{q}}$ = 400, 0^+ = 5 meV, and N_λ = 9 and 10 for $CsPbI_3$ and PbI_3^-, respectively, i.e., without the imaginary phonon bands. The blue crosses and the pink dot lines in Fig. 4, in variance, represent the set of data with a reduced number of **q** points ($N_{\mathbf{q}}$ = 375) and an increased damping parameter (0^+ = 10 meV), respectively. Since the plots are almost identical for both $CsPbI_3$ and PbI_3^-, I concluded that convergence was achieved when $N_{\mathbf{q}}$ = 400 and 0^+ = 5 meV. To validate my choice of excluding the imaginary phonon bands, the calculated lifetimes including the imaginary phonon bands are also shown in Fig. 4-8 (red dashed lines). The phonon frequencies were temporarily set to zero. We can see that the imaginary phonon bands have also no effect on my discussion concerning the trend in the carrier lifetimes.

From Fig. 4-8 (a), we can clearly see that the lifetime of the holes in the VB begins to increase at $E–E_{VBM}$ ~ –0.9 eV for all calculation settings. Because of the enhancement in the lifetime, the relaxation of the hot holes from the lower valence bands is suppressed at $E–E_{VBM}$ ~ –0.9 eV. Such expected dynamics agree exactly with the TAS measurements, which show slow hot-hole cooling in VB2 at ~0.9 eV below the VBM.[35] My calculations reproduce the experimental findings concerning $CH_3NH_3PbI_3$ perfectly.

The long-lifetime region shown in Fig. 4-8(a) is consistent with the small DOS region shown in Fig. 4-6(b). This is understandable, because a smaller DOS offers a reduced number of carrier relaxation paths leading to longer carrier lifetimes, and vice versa. A similar strong

relationship between the DOS and the e-ph scattering rate has been also reported for other semiconductors, e.g., silicon.[45–47] As expected, a longer-hole-lifetime region was not observed in PbI_3^- because it does not have a small DOS in the valence band. This correspondence between the electronic character and the carrier lifetime confirms my proposal that the hot-hole cooling mechanism can be explained mainly in terms of the characteristic electronic structure.

To discuss the carrier decay paths in detail, I examined the generalized Eliashberg functions of the Fan contribution

$$\mathrm{g}F^2_{n\boldsymbol{k}}(\omega)\big|_{\mathrm{Fan}} = \frac{1}{N_q}\sum_{n'\mathbf{q}\lambda}\frac{\left|g^{\mathbf{q}\lambda}_{nn'\mathbf{k}}\right|^2}{\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}+\mathbf{q}}}\left[2N_{\mathbf{q}\lambda}(T)+1\right]\delta\left(\omega-\omega_{\mathbf{q}\lambda}\right) \qquad (4.28)$$

This function revealed which phonon modes contributed to the carrier relaxation.

Fig. 4-9(a) shows the generalized Eliashberg functions of the two valence states shown in the inset band plot. In both plots, the highest peaks (b) were in the range 130 to 140 cm^{-1}. The hot holes are mainly thermalized through the phonon modes in this energy range. These phonon modes are longitudinal optical (LO) modes of the iodide ion vibration, which is the highest phonon band. The atomic motion of a phonon at (λ, **q**) = (15, (0, 0, 0.5)), which lies in the LO phonon band, is shown in Fig. 4-9 (b). It was also shown that the strength of peak (b) in the A state was lower than that in the B state, while the strength of the broad peak (c) ranging from 90 to 120 cm^{-1} increased at variance. The atomic motion of one of the phonon

modes in peak (c), which occurred at (λ, **q**) = (14, (0.25, 0.25, 0.25)), is also shown in Fig. 4-9(c). In this mode, not only the iodide ions but also the lead ions are active.

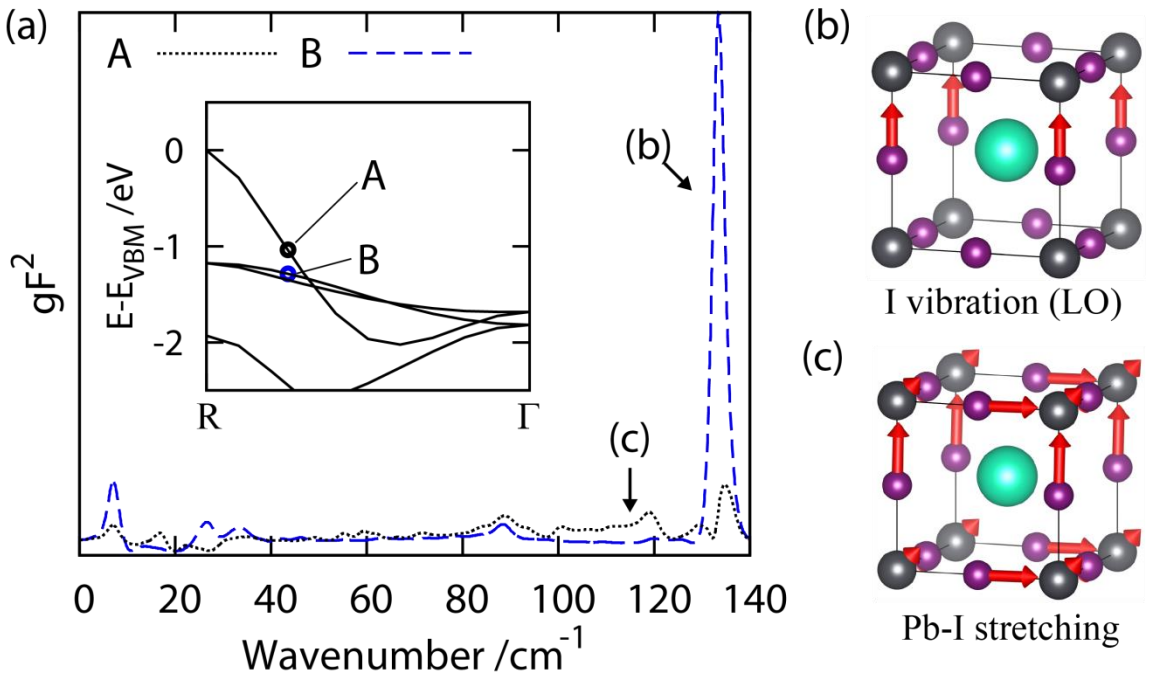


Figure 4-9. (a) Generalized Eliashberg functions of $CsPbI_3$ for the A and B valence states. Their positions are indicated in the inset band plot. (b), (c) Atomic motions of the phonon modes for (λ, **q**) = (15, (0, 0, 0.5)) and (λ, **q**) = (14, (0.25, 0.25, 0.25)), respectively. The black, purple, and green atoms denote Pb, I, and Cs, respectively.

In general, a single electronic state strongly couples with the phonon modes in which the atoms comprising the orbital are active. Thus, the strong e-ph interaction with the iodide ion vibrations can be ascribed to the fact that the valence band composition is mainly from I 5*p*

orbital. A similar explanation holds for the pronounced peak (c) in the A state as the highly dispersed valence band, where the A state lies, has a Pb 6*s* orbital contribution in addition to the I 5*p* orbital character as shown in our previous paper.[38]

Despite the enhancement of the peak (c) strength in the A state, it does not compensate for the reduction in the peak (b), and the overall e-ph coupling is weaker than that of the B state. This is ascribed to the fact that, at variance with the B states, the A state lying in the small DOS region has a reduced number of relaxation paths. This result clearly confirms my prediction that the electronic structure is the most relevant factor responsible for the e-ph lifetimes in lead halide perovskite materials.

As shown in Fig. 4-9, the vibrational modes coupled with the valence states of $CsPbI_3$ are ascribed to the motions of only I and Pb ions, not of Cs. This indicates that the replacement of the perovskite A-site cation does not impact on the mechanism of decay. Moreover, as shown in Fig. 4-6, the electronic structures of $CsPbI_3$ and $CH_3NH_3PbI_3$ are almost identical and the characteristically small DOS in the valence band is present in both $CsPbI_3$ and $CH_3NH_3PbI_3$. Therefore, my results about $CsPbI_3$ are easily extended to explain the experimental data concerning $CH_3NH_3PbI_3$.

The difference in the phonon properties of $CsPbI_3$ and $CH_3NH_3PbI_3$ is worth mentioning. In previous experimental and theoretical studies on the Raman spectrum of $CH_3NH_3PbI_3$, it was shown that the I–Pb–I bending and Pb–I stretching modes are coupled with librations of

$CH_3NH_3PbI_3$ in the wave-number range 60 to 100 cm^{-1}.[48] I predict that such coupled vibrations would still work as the relaxation paths of holes as long as the iodide ions are active. The torsional modes of $CH_3NH_3^+$ in the higher range 200–390 cm^{-1} are also coupled with the motion of the iodide ions through hydrogen bonds,[48] which may provide additional relaxation paths. However, the difference of wave-number range is assumed not to have a marked impact on the relaxation mechanism since the phonon energy scale is small compared with the energy scale for the electrons.

4-3-4. Conclusion

In conclusion, motivated by the experimental reports of slow hot-hole cooling from TAS measurements, I investigated the carrier e-ph lifetimes of lead iodide perovskites using first-principles calculations. My results show that holes in $CsPbI_3$ have longer lifetimes, particularly in the valence band region up to energy of 0.9 eV below the top of the valence band, where the DOS is characteristically small. On the other hand, such long hole lifetimes were not observed in PbI_3^-, which has a sharp DOS peak at the top of the valence band. Comparing these two results, I propose the origin of the slow hot-hole cooling to be the small DOS in the valence band, which is responsible for the reduced number of relaxation paths. My analysis of the generalized Eliashberg functions revealed that the holes decay via the motion of I and Pb ions, and not the motion of the A-site cation. I also predict that different

A-site cations do not impact on the carrier decay mechanism. A small DOS is indeed observed in the valence band of $CH_3NH_3PbI_3$, and thus the description of the fully inorganic halide perovskite $CsPbI_3$ can be easily extended to its organic–inorganic counterpart, $CH_3NH_3PbI_3$.

Bibliography

(1) Cannuccia, E.; Marini, A. *arXiv* **2013**, *1304,0072*.

(2) Baroni, S.; de Gironcoli, S.; Dal Corso, A. *Rev. Mod. Phys.* **2001**, *73*, 515.

(3) Cannuccia, E.; Marini, A. *Phys. Rev. Lett.* **2011**, *107*, 255501.

(4) As, D. J. *Microelectronics J.* **2009**, *40*, 204.

(5) Novikov, S. V; Stanton, N. M.; Campion, R. P.; Morris, R. D.; Geen, H. L.; Foxon, C. T.; Kent, a J. *Semicond. Sci. Technol.* **2008**, *23*, 015018.

(6) Ambacher, O.; Majewski, J.; Miskys, C.; Link, A.; Hermann, M.; Eickhoff, M.; Stutzmann, M.; Bernardini, F.; Fiorentini, V.; Tilak, V.; Schaff, B.; Eastman, L. F. *J. Phys. Condens. Matter* **2002**, *14*, 3399.

(7) Tschumak, E.; Granzner, R.; Lindner, J. K. N.; Schwierz, F.; Lischka, K.; Nagasawa, H.; Abe, M.; As, D. J. *Appl. Phys. Lett.* **2010**, *96*, 253501.

(8) Mietze, C.; Landmann, M.; Rauls, E.; Machhadani, H.; Sakr, S.; Tchernycheva, M.; Julien, F. H.; Schmidt, W. G.; Lischka, K.; As, D. J. *Phys. Rev. B* **2011**, *83*, 195301.

(9) Feneberg, M.; Röppischer, M.; Cobet, C.; Esser, N.; Schörmann, J.; Schupp, T.; As, D. J.; Hörich, F.; Bläsing, J.; Krost, A.; Goldhahn, R. *Phys. Rev. B* **2012**, *85*, 155207.

(10) Edmonds, K. W.; Novikov, S. V.; Sawicki, M.; Campion, R. P.; Staddon, C. R.; Giddings, a. D.; Zhao, L. X.; Wang, K. Y.; Dietl, T.; Foxon, C. T.; Gallagher, B. L. *Appl. Phys. Lett.* **2005**, *86*, 152114.

(11) Zainal, N.; Novikov, S. V.; Mellor, C. J.; Foxon, C. T.; Kent, a. J. *Appl. Phys. Lett.* **2010**, *97*, 112102.

(12) Mietze, C.; Lischka, K.; As, D. J. *Phys. Status Solidi* **2012**, *209*, 439.

(13) As, D. J.; Frey, T.; Schikora, D.; Lischka, K.; Cimalla, V.; Pezoldt, J.; Goldhahn, R.; Kaiser, S.; Gebhardt, W. *Appl. Phys. Lett.* **2000**, *76*, 1686.

(14) As, D. J.; Mietze, C. *Phys. Status Solidi* **2013**, *210*, 474.

(15) Benedict, L.; Shirley, E. *Phys. Rev. B* **1999**, *59*, 5441.

(16) Onida, G.; Reining, L.; Rubio, A. *Rev. Mod. Phys.* **2002**, *74*, 601.

(17) Aryasetiawan, F.; Gunnarsson, O. *Reports Prog. Phys.* **1998**, *61*, 237.

(18) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207.

(19) De Carvalho, L. C.; Schleife, A.; Bechstedt, F. *Phys. Rev. B* **2011**, *84*, 195105.

(20) Botti, S.; Marques, M. a. L. *Phys. Rev. Lett.* **2013**, *110*, 226404.

(21) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. *J. Phys. Condens. Matter* **2009**, *21*, 395502.

(22) Marini, A.; Hogan, C.; Grüning, M.; Varsano, D. *Comput. Phys. Commun.* **2009**, *180*, 1392.

(23) Del Sole, R.; Girlanda, R. *Phys. Rev. B* **1996**, *54*, 14376.

(24) Marini, A.; Del Sole, R. *Phys. Rev. Lett.* **2003**, *91*, 176402.

(25) Armiento, R.; Mattsson, a. *Phys. Rev. B* **2005**, *72*, 085108.

(26) Marini, A. *Phys. Rev. Lett.* **2008**, *101*, 106405.

(27) Logothetidis, S.; Petalas, J.; Cardona, M.; Moustakas, T. *Phys. Rev. B* **1994**, *50*, 18017.

(28) Kojima, A.; Teshima, K.; Shirai, Y.; Miyasaka, T. *J. Am. Chem. Soc.* **2009**, *131*, 6050.

(29) Burschka, J.; Pellet, N.; Moon, S.-J.; Humphry-Baker, R.; Gao, P.; Nazeeruddin, M. K.; Grätzel, M. *Nature* **2013**, *499*, 316.

(30) Liu, M.; Johnston, M. B.; Snaith, H. J. *Nature* **2013**, *501*, 395.

(31) Jeon, N. J.; Noh, J. H.; Kim, Y. C.; Yang, W. S.; Ryu, S.; Seok, S. Il. *Nat. Mater.* **2014**, *13*, 897.

(32) Grätzel, M.; Janssen, R. a J.; Mitzi, D. B.; Sargent, E. H. *Nature* **2012**, *488*, 304.

(33) Lee, M. M.; Teuscher, J.; Miyasaka, T.; Murakami, T. N.; Snaith, H. J. *Science* **2012**, *338*, 643.

(34) Etgar, L.; Gao, P.; Xue, Z.; Peng, Q.; Chandiran, A. K.; Liu, B.; Nazeeruddin, M. K.; Grätzel, M. *J. Am. Chem. Soc.* **2012**, *134*, 17396.

(35) Xing, G.; Mathews, N.; Sun, S.; Lim, S. S.; Lam, Y. M.; Grätzel, M.; Mhaisalkar, S.; Sum, T. C. *Science* **2013**, *342*, 344.

(36) Stranks, S. D.; Eperon, G. E.; Grancini, G.; Menelaou, C.; Alcocer, M. J. P.; Leijtens, T.; Herz, L. M.; Petrozza, A.; Snaith, H. J. *Science* **2013**, *342*, 341.

(37) Giorgi, G.; Fujisawa, J.; Segawa, H.; Yamashita, K. *J. Phys. Chem. Lett.* **2013**, *4*, 4213.

(38) Giorgi, G.; Fujisawa, J.-I.; Segawa, H.; Yamashita, K. *J. Phys. Chem. C* **2014**, *118*, 12176.

(39) Brivio, F.; Butler, K. T.; Walsh, A.; van Schilfgaarde, M. *Phys. Rev. B* **2014**, *89*, 155204.

(40) Ponseca, C. S.; Savenije, T. J.; Abdellah, M.; Zheng, K.; Yartsev, A.; Pascher, T.; Harlang, T.; Chabera, P.; Pullerits, T.; Stepanov, A.; Wolf, J.; Sundström, V. *J. Am. Chem. Soc.* **2014**, *136*, 5189.

(41) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048.

(42) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188.

(43) Huang, L.; Lambrecht, W. R. L. *Phys. Rev. B* **2013**, *88*, 165203.

(44) Stoumpos, C. C.; Malliakas, C. D.; Kanatzidis, M. G. *Inorg. Chem.* **2013**, *52*, 9019.

(45) Restrepo, O. D.; Varga, K.; Pantelides, S. T. *Appl. Phys. Lett.* **2009**, *94*, 212103.

(46) Bernardi, M.; Vigil-Fowler, D.; Lischner, J.; Neaton, J. B.; Louie, S. G. *Phys. Rev. Lett.* **2014**, *112*, 257402.

(47) Lautenschlager, P.; Allen, P.; Cardona, M. *Phys. Rev. B* **1986**, *33*, 5501.

(48) Quarti, C.; Grancini, G.; Mosconi, E.; Bruno, P.; Ball, J. M.; Lee, M. M.; Snaith, H. J.; Petrozza, A.; Angelis, F. De. *J. Phys. Chem. Lett.* **2014**, *5*, 279.

Chapter 5.
Concluding Remark

In this thesis, three semiconductors with enhanced photoenergy conversion capability were investigated using first-principles calculations using DFT and MBPT. The physical properties related to photoabsorption and carrier diffusion, i.e., bandgap, band bending, and carrier lifetime, were extensively discussed.

For overall water-splitting photocatalyst GaN:ZnO, I focused on the bandgap and band bending. The mechanism of bandgap reduction by chemical alloying of GaN and ZnO was investigated from two different approaches: slab model calculations to consider band edge positions and bulk model calculations to apply accurate but computationally demanding MBPT scheme. As a result, it was discovered that the bandgap reduction in GaN-rich alloys is ascribed to both VBM raising and CBM lowering. In particular, I found that the VBM position is governed by the local atomic configuration, i.e., the number and orientation of Zn-N bonds, while that of the CBM is not. Comparing computed absorption spectra with experiment, it is suggested that ZnO-rich solid solutions have larger number of Zn-N bonds than GaN-rich ones, which could be attributed to the low-temperature synthesis. Furthermore,

my result offered the possibility that downward band bending, which is opposite to upward bending usually occurred in n-type semiconductors, emerges at the surface of GaN-rich alloys.

Next, I focused on the effect of e-ph coupling in MBPT scheme to discuss carrier lifetimes. At first, this scheme was applied to *zb*-GaN. Collecting all MBPT schemes of GW approximation, BSE, and e-ph interaction, complete agreement of absorption spectrum with experiment was obtained. This result clearly demonstrated the importance of e-ph interaction on the electronic structure. Finally, the lifetime of hot carriers was investigated focusing on $APbI_3$, which is a promising material for perovskite solar cell. My result showed that $CsPbI_3$ has characteristic long lifetimes of the holes. The mechanism of such long lifetime was explained from the electronic point of view, i.e., characteristic small DOS. Thus such long lifetime is common trend in the other $APbI_3$ as far as small DOS is presented. It explains experimentally reported slow hot-hole cooling in $CH_3NH_3PbI_3$.

My works have demonstrated that the state-of-the-art first-principles calculations contribute the development of photoenergy conversion systems through quantitatively reliable analyses of bandgap, band bending, and carrier lifetimes. As a future outlook, impact of trap sites on recombination lifetime τ^{rec} must be addressed. Moreover, band bending in more realistic situation, i.e., solid-electrolyte interface, has to be considered for water-splitting photocatalyst. Such studies will complete understanding of photoenergy conversion process.

Publications stemming from this work

1. H. Kawai, K. Yamashita, E. Cannuccia, A. Marini, "Electron-electron and electron-phonon correlation effects on the finite-temperature electronic and optical properties of zinc-blende GaN", *Phys. Rev. B* **89**, 085202 (2014).
2. H. Kawai, G. Giorgi, A. Marini, K. Yamashita, "The mechanism of slow hot-hole cooling in lead-iodide perovskite: First-principle calculations on electron-phonon carrier lifetimes", submitted.
3. H. Kawai, G. Giorgi, and K. Yamashita, "First-principles study on photocatalytic properties of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$: Band-edge character and band bending of ZnO-doped GaN surface ", in preparation.
4. H. Kawai, G. Giorgi, M. Palummo, and K. Yamashita, "The role of local atomic configuration of $(Ga_{1-x}Zn_x)(N_{1-x}O_x)$ alloys on the visible light response: A combined GW + BSE first principle analysis", in preparation.