CHAPTER 1

INTRODUCTION

The interaction of light with matter has fascinated scientists over the ages. In 1670, Newton's first work as a Lucasian Professor at Cambridge was on optics. When he passed a beam of sunlight through a prism, Newton discovered that white light is not "pure" but is composed of a spectrum of light of all colors, from red to violet . In the nineteenth century, the linear theory of electricity and magnetism was developed involving some of the great mind of science: Faraday, Maxwell.

In the 1961, the second-harmonic generation experiment by Franken *et al.* marked the beginning of the field of nonlinear optics. Under the intense light of a laser, the properties of a material is modified such that the response of the material depends nonlinearly on the strength of the optical field. Nonlinear optics not only serves as a powerful tool to study various properties of materials but also finds wide applications such as in laser frequency conversion, optical communication, and optical switching. With solid state laser sources becoming more compact and new techniques developed to grow materials on the nanoscale, nonlinear optics will soon enter our everyday lives in the computers and telephones.

The quantitative understanding of the second-harmonic generation has involved several decades of work. The early developments of the theory include contributions from Bloembergen, Kleinman, B. Levine, and Aspnes within the first ten years after Franken's first experiment. Over the last decade, with the help of larger and faster computers, it has become possible to calculate the second-harmonic susceptibilities from *ab initio* methods. In particular, Z. H. Levine and coworkers succeeded in predicting the second-harmonic susceptibilities within 20% agreement compared to experiment.

In this work, Z. H. Levine's method is extended to study a class of semiconductors with large band gaps, including four polytypes of SiC (2H, 4H, 6H, and 3C), BN, AlN, and GaN. This work also formally completes Levine's theory by including the the local-field corrections that arises from the nonlinearly induced microscopic charge density.

1.1 Background on the second-harmonic susceptibility

Light passing through a material induces a polarization $P(t)$ which is conventionally written in a power series in the electric field of light $E(t)$. The coefficients in the power series are the linear and nonlinear susceptibilities $\chi^{(i)}$ $(i = 1, 2, \cdots)$ that describe how the materials interact with light

$$P(t) = \chi^{(1)} E(t) + \frac{1}{2}\chi^{(2)} E(t)E(t) + \cdots. \tag{1.1}$$

The linear susceptibility $\chi^{(1)}$ is of order unity; the second-order susceptibility $\chi^{(2)}$ is small, usually on the order of pm/V $= (4\pi/30) \times 10^{-9}$ esu. The nonlinear effects, including the second-, third-, and higher-order effects, can be observed only under intense light such as laser. The second-order susceptibility, in particular, is a third-order tensor and vanishes for materials with inversion symmetry.

The second-harmonic generation is a particular case of the second-order nonlinear optical responses.* First observed by Franken *et al.* in 1961,[2] the second-harmonic generation is related to the process in which the induced polarization has twice the frequency of the incoming light,

$$P^{(2)}(2\omega) = \frac{1}{2}\chi^{(2)}(-2\omega; \omega, \omega)E(\omega)E(\omega), \tag{1.2}$$

where $\chi^{(2)}(-2\omega; \omega, \omega)$ is the second-harmonic susceptibility. Another frequently used definition[3] of the second-harmonic susceptibility d differs from $\chi^{(2)}$ by a factor of 2,

$$P^{(2)}(2\omega) = d(-2\omega; \omega, \omega)E(\omega)E(\omega). \tag{1.2$'$}$$

* Other second-order responses include the sum- and difference-frequency generations. See Ref. 1 for more details.

An important application of the second-harmonic generation is in the efficient frequency-doubling devices. Up to date, there are no solid state lasers in the ultra-violet region. The first blue laser diode was demonstrated only earlier this year.[4] An alternative method to obtain blue and ultraviolet laser light is through the use of nonlinear crystals to convert the red or infra-red lasers to their second-harmonic frequencies. Some of the most frequently used nonlinear crystals include ammonium dihydrogen phosphate, $NH_4H_2PO_4$ (ADP), potassium dihydrogen phosate, KH_2PO_4 (KDP), and lithium niobate, $LiNbO_3$.[5]

A stringent condition is placed on the crystals that can be used for efficient frequency doubling devices – the phase-matching condition.[5] This condition requires that the refractive index $n(\omega)$ of the crystal satisfies

$$n(2\omega) = n(\omega), \tag{1.3}$$

so that the generated second-harmonic light signal (with 2ω frequency) remains phase-coherent with the pump light throughout the crystal. Under such a condition, the intensity of the second-harmonic light grows with the length of the light path.[5] The most commonly used technique to achieve phase-matching is through the use of the birefringence displayed by many crystals.[5] Not all crystals display birefringence, in particular, crystals having the cubic symmetry are optically isotropic and thus are not phase-matchable.

The main goal of our study is to construct a reliable theoretical method which can predict the second-harmonic susceptibilities of semiconductors and insulators from first-principles. The materials of our current study, e.g., the group III-V and IV-IV compounds, have simple structures such as the zincblende or wurtzite structures. Only after testing our method on these simple materials can we apply the method to more complicated structures with larger system sizes to further test our method and predict the second-harmonic susceptibilities for new materials.

1.2 The classical model for the second-harmonic susceptibility

A classical model provides a simple way to understand the microscopic origin and some qualitative features of the second-harmonic susceptibility. More sophisticated versions of the classical or semi-classical models, including the bond-charge model[6] and the two-band model,[7] give predictions which are in factor-of-two agreement with experiment for many materials. In this section, we derive the classical model for the second-harmonic susceptibility using the harmonic oscillator with a small cubic term. The most successful of the semi-classical models – the bond-charge model – will be briefly described. We also discuss the bond-additivity, which is an underlying assumption of all the classical models.[8]

Consider a classical particle with charge q and mass m in a 1-dimensional potential which is mostly harmonic $m\omega_0^2 x^2/2$ but with a small cubic term $m\beta x^3/3$ (see Figure 1.1). The equation of motion for the particle in an external electric field $Ee^{-i\omega t}$ is given by Newton's second law

$$m\frac{d^2x}{dt^2} + m\omega_0^2 x + m\beta x^2 = qEe^{-i\omega t}. \tag{1.4}$$

The solution to this equation up to the second-order in E is

$$x = \frac{q}{m}\frac{1}{\omega_0^2 - \omega^2}Ee^{-i\omega t} + \frac{\beta q^2}{m^2}\frac{1}{[\omega_0^2 - (2\omega)^2](\omega_0^2 - \omega^2)^2}E^2 e^{-2i\omega t}. \tag{1.5}$$

If there are N such oscillators, then from the total polarization $P = Nqx$ and Eq. (1.1), we obtain the linear and the second-harmonic susceptibilities

$$\begin{aligned}\chi^{(1)} &= \frac{Nq^2}{m}\frac{1}{\omega_0^2 - \omega^2},\\ \chi^{(2)} &= \frac{N\beta q^3}{m^2}\frac{1}{[\omega_0^2 - (2\omega)^2](\omega_0^2 - \omega^2)^2}.\end{aligned} \tag{1.6}$$

Eq. (1.6) shows that the second-harmonic susceptibility originates from the cubic part of the potential described by β, and diverges at $\omega = \omega_0/2$ and ω_0. In insulating materials, ω_0 is associated with the band gap.

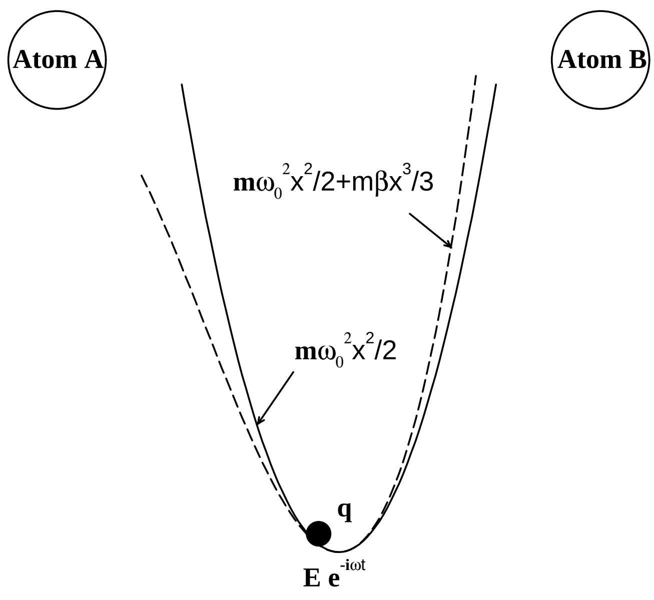

FIGURE 1.1 Illustration of the second-harmonic susceptibility from a classical model. A particle, with charge q, moves in a potential between two atoms. If the potential is harmonic, $m\omega_0^2/2$, the particle responds only linearly to an external field $Ee^{-i\omega t}$. Adding a small cubic term to the potential, $m\beta x^3/3$, for example when the two atoms are of different species, yields a second-order nonlinear response of the particle.

The bond-charge model for the second-harmonic susceptibility was developed by Levine in the 60-70's and was successful in reaching a factor-of-two agreement with experiment for over 100 materials.[6] A point charge is associated with each bond in a crystal and is polarizable only in the bonding direction. By fitting to data on linear response, the band gap can be obtained with is related to ω_0. The potential on the point charge from the core electrons is approximated by the Thomas-Fermi potential. The differences between the ionicities and the core-radii of the ions give rise to an asymmetric potential from which a cubic term is extracted. With the bond-charge, the band gap, and the cubic potential, the microscopic second-order polarizability of each bond can be computed.[6] The above is an over-simplified description of the bond-charge model. See Ref. 6 for more details.

To obtain the macroscopic second-harmonic susceptibilities of crystals from the microscopic polarizability of each individual bond, one needs another assumption – the assumption of the bond-additivity.[8] Under this assumption, the bond charge polarizes only in the bonding direction (see Figure 1.2) and the interaction between the bonds is neglected. By a tensorial sum of the microscopic polarizabilities of all the bonds in the crystal, the macroscopic second-harmonic susceptibility can be obtained.[8] The bond-additivity is the underlying assumption of all the classical and semi-classical models.

For a tetrahedrally coordinated binary compound with a single type of bond, the bond-additivity gives a simple relation between the different components of the macroscopic second-harmonic susceptibility coefficients.[8] Let $\chi^{(2)}(micro)$ be the second-order polarizability of each bond, and θ_i be the angle between the bond and the i-axis (see Figure 1.2). The ijk component of the macroscopic second-harmonic

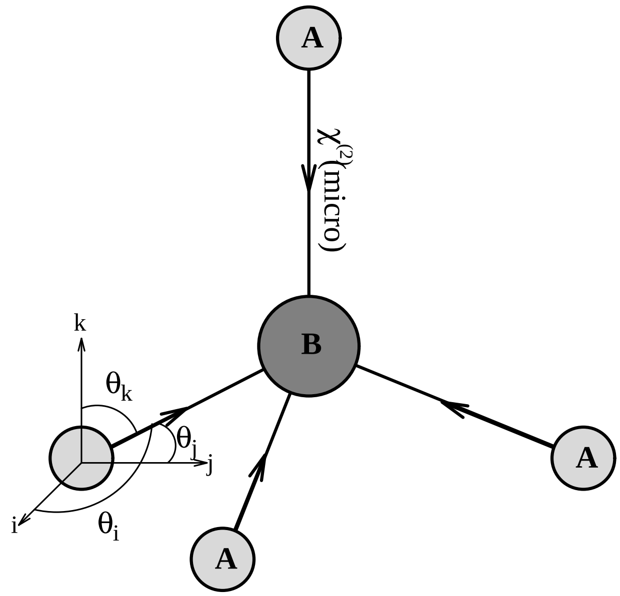

FIGURE 1.2 Illustration of using the bond-additivity to calculate the macroscopic second-harmonic susceptibility $\chi^{(2)}_{ijk}(macro)$ from the microscopic second-order polarizability $\chi^{(2)}(micro)$. The angles between the directions of the bonds and axes are denoted by θ_i, θ_j, and θ_k. The shaded circles represent the atoms. To obtain $\chi^{(2)}_{ijk}(macro)$, one uses a tensorial sum of $\chi^{(2)}(micro)$ from all the bonds, which is described by Eq. (1.7) in the text.

susceptibility $\chi^{(2)}_{ijk}(macro)$ for a crystal with unit cell volume Ω_0 is given by

$$\chi^{(2)}_{ijk}(macro) = \frac{1}{\Omega_0} \sum_{bonds} \chi^{(2)}(micro) \cos\theta_i \cos\theta_j \cos\theta_k, \tag{1.7}$$

where the sum is over the bonds in the unit cell.[8] For the wurtzite structure, in particular, there are two independent components of $\chi^{(2)}(macro)$ from the symmetries of the crystal, $\chi^{(2)(w)}_{xxz}(macro)$ and $\chi^{(2)(w)}_{zzz}(macro)$.[9] With only one type of bond in the wurtzite structure and hence a single $\chi^{(2)}(micro)$, the sum in Eq. (1.7) can be evaluated to give[8]

$$\begin{aligned}\chi^{(2)(w)}_{xxz}(macro) &= -\frac{4}{9}\frac{\chi^{(2)}(micro)}{\Omega_0},\\ \chi^{(2)(w)}_{zzz}(macro) &= \frac{8}{9}\frac{\chi^{(2)}(micro)}{\Omega_0}.\end{aligned} \tag{1.8}$$

For some compounds, both the wurtzite and the zincblende structures exist. Assuming the bonds are the same in both structures, the $\chi^{(2)(z)}_{xyz}(macro)$ of the zincblende structure can be obtained analogously

$$\chi^{(2)(z)}_{xyz}(macro) = -\frac{2}{3\sqrt{3}}\frac{\chi^{(2)}(micro)}{(\Omega_0/2)}, \tag{1.9}$$

where the unit cell volume of the zincblende structure is half that of the wurtzite structure. Comparing Eqs. (1.8) and (1.9) yields

$$\chi^{(2)(w)}_{xxz}(macro) = -\frac{1}{2}\chi^{(2)(w)}_{zzz}(macro) = -\frac{1}{\sqrt{3}}\chi^{(2)(z)}_{xyz}(macro). \tag{1.10}$$

Experimental measurements of the ratio $\chi^{(2)(w)}_{xxz}(macro)/\chi^{(2)(w)}_{zzz}(macro)$ on ZnS, CdS, CdSe, and GaN are to be -0.53, -0.57, -0.55, and -0.50, respectively,[6] in close agreement with the first equality of (1.10) for which the rule of thumb prediction is negative one-half. However, measurements on BeO, and AlN give the ratio to be $+0.73$ and ≤ 0.05, which contradict with Eq. (1.10). As will be shown in Secs. (3.2) and (3.3), our calculations on 4H-SiC, 6H-SiC, and BN are consistent with Eq. (1.10); but our calculations on 2H-SiC, AlN, and GaN are not.

The classical model gives a clear picture on the origin of the second-harmonic susceptibility and provides a simple method to compute it. However, as shown in the tests of the bond-additivity, the results from classical models are, in a few cases, qualitatively different from experiment. For a more reliable method to calculate the second-harmonic susceptibilities, we use the quantum mechanical approach.

The methodology is much more complicated and no clear picture has yet emerged. However, the results we obtain are all in quantitative agreement with experiment (see Table 4.3).

1.3 The local-density approximation

Quantum mechanics provides a complete set of equations for the electrons which interact with each other through the long-range Coulomb potential. However, it is impossible to solve these equations exactly in systems with even a few tens of electrons, because all the equations are coupled. In solids where there are 10^{23} electrons, approximations have to be made before any calculation can take place. In this section, we give a very brief introduction to the density functional theory which maps the interacting many-electron problem to a single-electron mean-field problem.[10] All the unknowns of the mean-field potentials are included in the exchange-correlation potential which, in practice, is approximated by a local potential depending on only the local electronic density, and hence local-density approximation (LDA).[11]

The density functional theory,[10] introduced by Hohenberg and Kohn in 1964, states that the ground state electronic density corresponds to the density $n(\vec{r})$ that minimizes the energy functional $E[n]$, which in atomic units is written as

$$E[n] = T[n] + \frac{1}{2}\int d^3r d^3r' \frac{n(\vec{r})n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} + E_{xc}[n] + \int d^3r V_{ext}(\vec{r})n(\vec{r}), \qquad (1.11)$$

where $T[n]$ is the kinetic energy of the noninteracting electrons, the second term on the right hand side is the classical Coulomb (or Hartree) interaction, $E_{xc}[n]$ is the exchange-correlation energy functional. The ionic potentials and external fields are included in $V_{ext}(\vec{r})$. Minimizing $E[n]$ subject to the constraint that the number of the electrons is conserved, i.e.,

$$\int d^3 r n(\vec{r}) = N, \tag{1.12}$$

yields the Euler equation with the chemical potential μ

$$\frac{\delta T[n]}{\delta n(\vec{r})} + \int d^3 r' \frac{n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} + v_{xc}(\vec{r}) + V_{ext}(\vec{r}) - \mu = 0, \tag{1.13}$$

where $v_{xc}(\vec{r}) \equiv \delta E_{xc}[n]/\delta n(\vec{r})$, is the exchange-correlation potential.

In 1965, Kohn and Sham extended the density functional theory to make calculations possible.[11] They showed that Eq. (1.13) is satisfied by self-consistently solving the set of equations,[11] which in the non-spin-polarized case are

$$\begin{aligned} H_{KS}\psi_i(\vec{r}) &= \left[-\frac{\nabla^2}{2} + \int d^3 r \frac{n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} + v_{xc}(\vec{r}) + V_{ext}(\vec{r}) \right] \psi_i(\vec{r}) = \epsilon_i \psi_i(\vec{r}), \\ n(\vec{r}) &= 2 \sum_{i=1}^{N/2} |\psi_i(\vec{r})|^2, \end{aligned} \tag{1.14}$$

where $\{\psi_i\}$ are orthonormal eigenfunctions and $\{\epsilon_i\}$ are the associated eigenvalues of the Kohn-Sham Hamiltonian H_{KS}. The sum in Eq. (1.14) extends over $N/2$ lowest eigenstates. Note that the functional form of v_{xc} is still unspecified.

The local-density approximation (LDA) assumes that the exchange-correlation potential is purely local and can be written as,[11]

$$E_{xc}[n] = \int d^3 r \varepsilon_{xc}(n(\vec{r})) n(\vec{r}), \tag{1.15}$$

where $\varepsilon_{xc}(n(\vec{r}))$ is the exchange-correlation energy per particle of a uniform electron gas of density $n(\vec{r})$. This assumption is valid when the density varies slowly in space. The functional form of $\varepsilon_{xc}(n)$ is given by several parameter fits. One of the fits is given by Wigner,[12] which, in terms of the mean distance between the electrons $r_s \equiv (3/4\pi n)^{1/3}$, is

$$\varepsilon_{xc}^{Wigner}(r_s) = -\frac{3}{4}\left(\frac{3}{2\pi}\right)^{2/3}\frac{1}{r_s} - \frac{0.44}{r_s + 7.79}. \tag{1.16}$$

The most frequently used parameters are obtained by fitting to the quantum Monte Carlo calculations by Ceperley and Alder on free electron gas.[13] The particular parameters we use in our calculations are given by Teter,[14]

$$\varepsilon_{xc}^{Teter}(r_s) = -\frac{a_0 + a_1 r_s + a_2 r_s^2 + a_3 r_s^3}{b_1 r_s + b_2 r_s^2 + b_3 r_s^3 + b_4 r_s^4}, \tag{1.17}$$

where

$$\begin{cases} a_0 = (3/4)(3/2\pi)^{2/3}, & b_1 = 1, \\ a_1 = 2.40875407, & b_2 = 4.91962865, \\ a_2 = 0.88642404, & b_3 = 1.34799453, \\ a_3 = 0.02600342, & b_4 = 0.03120453. \end{cases}$$

The density-functional theory implemented with LDA has been successful in studying the ground-state properties of atoms, molecules, simple and transition metals, semiconductors, surfaces, etc.[15] There have been consistent efforts to improve LDA with the "general gradient approximation" (GGA) which includes the dependence on the gradient of the density to the exchange-correlation functional.[16] Using GGA, however, has not yielded signifant improvements to LDA on the dielectric constants.[17] In the current study, we will use LDA as a starting point of our optical response calculations. Caution needs be taken in interpreting the eigenvalues ϵ_i given by the mean-field Kohn-Sham Hamiltonian as the quasiparticle energies.[18] We will discuss this point further in Sec. 1.5.

1.4 Review of the calculations of the second-harmonic susceptibility

In this section, we introduce the sum-over-states method to calculate the second-harmonic susceptibility. The historical development of the method is reviewed.

When light passes through an insulating crystal, its electric field drives the electrons to oscillate around their equilibrium positions. The collective motion of electrons can be described by an induced macroscopic current that follows the electric field. The macroscopic current can be computed microscopically using perturbation theory to calculate the contributions from all the valence bands and all $\vec{k}$-points. Calculation of $\chi^{(2)}$ involves a second-order perturbative calculation of the change in the density matrix which leads to a second-order induced current.[19] The expression for $\chi^{(2)}$ thus obtained can be broken into the ccv and vvc contributions ($c \to$ conduction, $v \to$ valence) corresponding to the virtual electron and virtual hole processes, respectively (see Figure 1.3).[20] Denoting the wavefunctions by $|n\vec{k}\rangle$, the eigenvalues by $\epsilon_{n\vec{k}}$, and the momentum operator by P, the static limits for the expressions of the ccv and vvc contributions are

$$\chi^{(2)} = \chi^{(2)}(ccv) + \chi^{(2)}(vvc),$$

$$\chi^{(2)}(ccv) = \sum_{v}\sum_{c}\sum_{c'}\int d^3k \;\; \mathrm{Im}\Big(\langle v\vec{k}|P|c\vec{k}\rangle\langle c\vec{k}|P|c'\vec{k}\rangle\langle c'\vec{k}|P|v\vec{k}\rangle\Big)$$
$$\times \frac{3(\epsilon_{c\vec{k}} - \epsilon_{c'\vec{k}})\Big[2(\epsilon_{v\vec{k}} - \epsilon_{c\vec{k}})^2 + 3(\epsilon_{v\vec{k}} - \epsilon_{c\vec{k}})(\epsilon_{v\vec{k}} - \epsilon_{c'\vec{k}}) + 2(\epsilon_{v\vec{k}} - \epsilon_{c'\vec{k}})^2\Big]}{(\epsilon_{v\vec{k}} - \epsilon_{c\vec{k}})^4(\epsilon_{v\vec{k}} - \epsilon_{c'\vec{k}})^4}$$

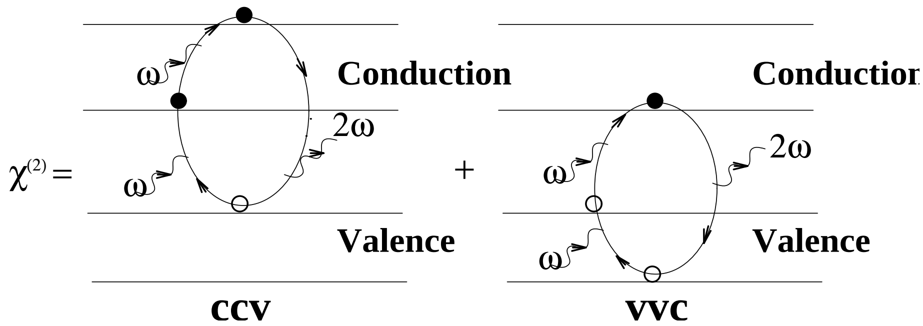

FIGURE 1.3 The *ccv* and *vvc* contributions to the second-harmonic susceptibility $\chi^{(2)}$. The *ccv* term corresponds to the virtual electron process in which a valence electron absorbs a photon with frequency ω and hops onto a conduction state. The electron absorbs a second photon and hops onto another conduction state before recombining with the hole and emitting a photon with frequency 2ω. The *vvc* term corresponds to the virtual hole process in which the hole absorbs a second photon and hops onto another valence state.

$$\chi^{(2)}(vvc) = \sum_{v}\sum_{v'}\sum_{c}\int d^3k \;\; \mathrm{Im}\Big(\langle v\vec{k}|P|v'\vec{k}\rangle\langle v'\vec{k}|P|c\vec{k}\rangle\langle c\vec{k}|P|v\vec{k}\rangle\Big)$$
$$\times \frac{3(\epsilon_{v'\vec{k}}-\epsilon_{v\vec{k}})\Big[2(\epsilon_{v\vec{k}}-\epsilon_{c\vec{k}})^2+3(\epsilon_{v\vec{k}}-\epsilon_{c\vec{k}})(\epsilon_{v'\vec{k}}-\epsilon_{c\vec{k}})+2(\epsilon_{v'\vec{k}}-\epsilon_{c\vec{k}})^2\Big]}{(\epsilon_{v\vec{k}}-\epsilon_{c\vec{k}})^4(\epsilon_{v'\vec{k}}-\epsilon_{c\vec{k}})^4}. \tag{1.18}$$

With the eigenvalues and wavefunctions determined, e.g., using LDA, the momentum matrix elements can be constructed and hence $\chi^{(2)}$ can be computed. Calculations based on the above expressions are categorized as using the sum-over-states method.

The first calculation of the second-harmonic susceptibility $\chi^{(2)}$ using the sum-over-states method dates back over twenty years. In 1975, Fong and Shen[21] used empirical pseudopotentials to calculate $\chi^{(2)}$ for GaAs, InAs, and InSb. Due to the inaccurate wave functions and hence the momentum matrix elements, the calculated values are an order of magnitude smaller than experiment (see Table 1.1).[21] In 1987, Moss, Sipe, and van Driel used empirical tight-binding band structures to obtain $\chi^{(2)}$ for GaP, GaAs, GaSb, and InAs which agree with experiment within a factor of four.[22] In 1991, Ghaharamani, Moss, and Sipe improved their technique by using a minimal-basis orthogonalized linear combination of Gaussian orbitals method and obtained better results (see Table 1.1).[23] In 1991, Ma, Chui, Kosowski, and Hsu calculated $\chi^{(2)}$ for GaAs using pseudopotentials and plane-waves.[24] The value they obtained is three times larger than experiment. In 1992, Huang and Ching used the self-consistent orthogonalized-linear-combination-of-atomic orbitals to calculate $\chi^{(2)}$ of GaAs.[25] They later extended the calculations to 15 semiconductors including group III-V, and II-VI compounds and reached a factor of two agreement with experiment.[26] Earlier in 1991, Levine and Allan calculated $\chi^{(2)}$ for AlP, AlAs, GaP, and GaAs using plane-waves and pseudopotentials.[27] They also introduced the "scissors operators" to correct the too small LDA band gap.[28] Their calculated $\chi^{(2)}$ with the scissors operators are in excellent agreement with experiment.[27] Levine, Wilkins, and co-workers later extended their method to successfully reproduce the experimental results on Se,[29] α-quartz,[29] and urea.[30] Table 1.1 lists the computed $d = \chi^{(2)}/2$ for GaAs using the different methods for comparison.

Recently, a new method was proposed by Dal Corso *et al.*, who use the "$2n + 1$" theorem of the perturbation theory to arrive at an alternative expression for $\chi^{(2)}$ which is simpler than Eq. (1.18).[31] As we will prove in Appendix C, the new expression, Eq. (C11), of the "$2n + 1$" method is analytically equivalent to the expression of the sum-over-states method. The "$2n + 1$" method is also easier to

TABLE 1.1 Comparison of the calculated second-harmonic susceptibility $d_{xyz} = \chi^{(2)}_{xyz}/2$ for GaAs using different methods. The calculations listed are done at the experimental lattice constant with the exception of the calculation by Dal Corso *et al.* which is carried out at the LDA determined lattice constant. From Ref. 27, d_{xyz} is 50% smaller at the LDA lattice constant than at the experimental lattice constant. Local-field corrections are included in the calculations by Levine *et al.*, which leads to an 8% reduction in d_{xyz}, and the calculation by Dal Corso *et al.* The calculation that best agrees with experiment is done by Levine *et al.* with the scissors operators to shift the too small LDA conduction band gap to the experimental gap.

Method	d_{xyz} (pm/V)
Empirical pseudopotential (Fong, 1976)[21]	8
Empirical tight-binding (Moss, 1987)[22]	201
Linear-combination-of-Gaussian-orbitals (Ghahramani, 1991)[23]	120
Plane-wave pseudopotential (Ma, 1991)[24]	400
Linear-combination-of-atomic-orbitals (Huang, 1992)[25]	120
Plane-wave pseudopotential – LDA (Levine, 1991)[27]	174
Plane-wave pseudopotential – Scissors (Levine, 1991)[27]	86
Plane-wave pseudopotential – "$2n+1$" (Dal Corso, 1996)[31]	103
Experiment[3]	81±5

evaluate numerically because the expression for $\chi^{(2)}$ has no explicit dependence on the conduction bands and hence avoids the diagonalization of the Hamiltonian matrix which is required in the sum-over-states method. The author feels that the "$2n+1$" method will eventually replace the traditional sum-over-states method and make possible the calculations of $\chi^{(2)}$ for larger systems and complex structures.

In our current study, the sum-over-states method developed by Levine *et al.* is extended to study the effects of the scissors operators which has been successful in the small to medium gap materials.[32] Since the two methods are formally equivalent, the conclusions drawn from the studies using the sum-over-states method provide a guide to future studies of $\chi^{(2)}$ with the "$2n + 1$" method.

1.5 The GW approximation and the scissors operators

The eigenvalues of the single-particle Kohn-Sham Hamiltonian do not correspond to the quasiparticle energies in general.[18] In semiconductors and insulators, for example, the band gap E_g given by LDA is always smaller than the experimental gap.[33] In Si, E_g^{LDA}=0.5 eV and E_g^{Exp}=1.17 eV. Here we introduce the GW approximation[34] (G: Green's function, W: screened Coulomb potential) which gives quasiparticle energies in good agreement with experiment. The effects of the GW calculations for optical response studies can, in a crude approximation, be captured by the "scissors operators" which shift the conduction band energies by a constant.[28]

The quasiparticle energies $\epsilon_{n\vec{k}}^{QP}$ and wavefunctions $\psi_{n\vec{k}}^{QP}(\vec{r})$ are obtained by solving the Dyson's equation

$$\left[-\frac{\nabla^2}{2}+\int d^3r'\frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}+V_{ext}(\vec{r})\right]\psi^{QP}_{n\vec{k}}(\vec{r})+\int d^3r'\Sigma(\vec{r},\vec{r}';\epsilon_{n\vec{k}})\psi^{QP}_{n\vec{k}}(\vec{r})$$
$$=\epsilon^{QP}_{n\vec{k}}\psi^{QP}_{n\vec{k}}(\vec{r}), \tag{1.19}$$

where the self-energy operator Σ replaces the exchange-correlation potential $v_{xc}(\vec{r})$ in the Kohn-Sham equation Eq. (1.14).[34,35] The self-energy operator Σ contains the exchange-correlation effects among the electrons. In general, the self-energy $\Sigma(\vec{r},\vec{r}\,';E)$ is nonlocal, energy-dependent, and nonhermitian. The real part of the eigenvalues $\epsilon^{QP}_{n\vec{k}}$ gives the energies of the quasiparticle and the imaginary part gives the lifetime. So far, there does not exist a systematic correction to the exchange-correlation potential $v_{xc}(\vec{r})$ for the self-energy Σ, because, unlike the Hartree-Fock theory, LDA does not have a known expansion in terms of the Feymann diagrams.

The GW approximation gives a practical and successful solution to constructing the self energy starting from LDA calculations. Let $G(\vec{r},\vec{r}\,';E)$ be the single-particle Green's function and $W(\vec{r},\vec{r}\,';E)$ be the dynamically screened Coulomb potential. The self energy of the GW approximation is given by[34,35]

$$\Sigma(\vec{r},\vec{r}\,';E)=i\int\frac{dE'}{2\pi}e^{-i\delta E'}G(\vec{r},\vec{r}\,';E'-E)W(\vec{r},\vec{r}\,';E'), \tag{1.20}$$

where $\delta=0^+$. The Green's function can be constructed from the eigenvalues $\epsilon^{LDA}_{n\vec{k}}$ and the corresponding eigenfunctions $\psi^{LDA}_{n\vec{k}}(\vec{r})$ of the LDA Hamiltonian

$$\Sigma(\vec{r},\vec{r}\,';E)=\sum_{n,\vec{k}}\frac{\psi^{LDA}_{n\vec{k}}(\vec{r})\psi^{LDA}_{n\vec{k}}(\vec{r}\,')^*}{E-\epsilon^{LDA}_{n\vec{k}}-i\delta}, \tag{1.21}$$

where $\delta=0^+$ for $\epsilon^{LDA}_{n\vec{k}}<\mu$ and $\delta=0^-$ for $\epsilon^{LDA}_{n\vec{k}}>\mu$; μ is the chemical potential. Let $W_{\vec{G},\vec{G}\,'}(\vec{q};E)$ be the fourier transform of the dynamically screened Coulomb potential,

$$W(\vec r,\vec r\,';E)=\sum_{\vec q,\vec G,\vec G\,'} e^{i(\vec q+\vec G)\cdot\vec r}W_{\vec G,\vec G\,'}(\vec q;E)e^{-i(\vec q+\vec G\,')\cdot\vec r\,'}. \tag{1.22}$$

We then have

$$W_{\vec G,\vec G\,'}(\vec q;E)=\frac{\varepsilon^{-1}_{\vec G,\vec G\,'}(\vec q,E)}{|\vec q+\vec G|^2}, \tag{1.23}$$

where the dielectric constant matrix $\varepsilon_{\vec G,\vec G\,'}(\vec q,E)$ can be computed from the random phase approximation[36]

$$\begin{aligned}\varepsilon_{\vec G,\vec G\,'}(\vec q,E)=\delta_{\vec G,\vec G\,'}-\frac{4\pi}{|\vec q+\vec G|^2}\sum_{n,m,\vec k}&\frac{f(\epsilon_{m\vec k+\vec q})-f(\epsilon_{n\vec k})}{\epsilon_{m\vec k+\vec q}-\epsilon_{n\vec k}+E}\\ &\times\langle n\vec k|e^{-i(\vec q+\vec G)\cdot\vec r}|m\vec k+\vec q\rangle\langle m\vec k+\vec q|e^{i(\vec q+\vec G\,')\cdot\vec r\,'}|n\vec k\rangle.\end{aligned} \tag{1.24}$$

in which f is the Fermi occupation number.

The GW calculations are an order of magnitude more intensive numerically than the LDA and have been carried out for only a handful of materials.[35,37,38,39,40,41] The band gaps of these GW calculations are very close to the experimental values, usually within 0.3 eV. Comparing the results of the GW and LDA calculations shows that the corresponding valence band wavefunctions have high overlap, over 95% as reported in Ref. 35. The largest change to the LDA eigenvalues is a rigid shift to the conduction band energies. These two observations – (1) the LDA wavefunctions remain unchanged, and (2) the quasiparticle energy spectrum can be approximated from the LDA calculation by a rigid shifted to the conduction band energies – led to the proposal of the scissors operators by Levine and Allan in 1989.[28]

Prior to the introduction of the scissors operators, the calculated dielectric constant ε_∞ using LDA are larger than experimental values by over 10%, e.g., in Si

ε_∞^{LDA}=12.7 and ε_∞^{Exp}=11.4.[42] Levine and Allan associated the overestimation of ε_∞ with the too small LDA band gap.[28] Indeed from Eq. (1.6) of the classical model, if ω_0, which is associated with the band gap, is too small, the dielectric constant $\varepsilon_\infty = 1 + 4\pi\chi^{(1)}$ becomes too large. Motivated by the results of the GW calculations,[35] Levine and Allan proposed the "scissors-operator" approximation which applies to the calculations of optical response coefficients.[28] As illustrated in Figure 1.4, the quasiparticle/experimental energy spectrum can be obtained by cutting the LDA energy spectrum between the gap with a pair of "scissors" and shifting the conduction bands by a constant Δ. In $\vec{k}$-space, the Hamiltonian corresponding to the scissors operators $H_{\vec{k}}^{sci}$ is equal to the LDA Hamiltonian $H_{\vec{k}}^{LDA}$ plus the energy shift Δ times the projection onto the conduction bands, i.e.,

$$H_{\vec{k}}^{sci} = H_{\vec{k}}^{LDA} + \Delta \sum_{n}^{cond} |n\vec{k}\rangle\langle n\vec{k}|. \tag{1.25}$$

By using the scissors operators approach, Levine and co-workers successfully reproduced several other optical response coefficients including the dielectric constants and their strain-induced effects, and the optical activity-coefficients, for over a dozen small to medium semiconductors.[32] In particular, for Si ε_∞^{sci}=10.94, which agrees with the experimental value to 5%. For the second-harmonic susceptibilities, the scissors operator has a larger effect. In GaAs, for example (see Table 1.1), the scissors operator leads to a 50% reduction in d and good agreement with experiment.

In the current study, we extend the optical response studies to wide gap semiconductors to further test the scissors operators. As we will show in Chapter 3, the optical response coefficients in the wide gap materials calculated with LDA are larger than experiment. However, in contrast to the small medium gap materials, the coefficients given by scissors operators are smaller than experiment. This

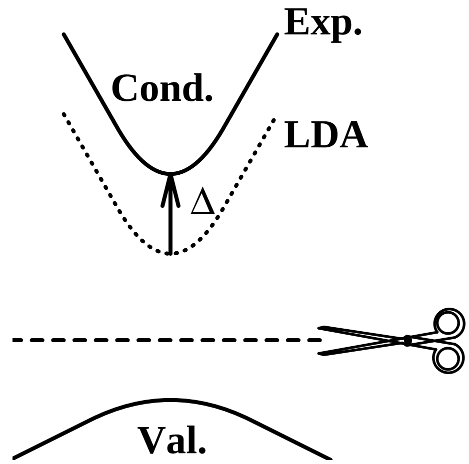

FIGURE 1.4 Illustration of the "scissors operator". In semiconductors and insulators, the band gaps given the LDA calculations are always smaller than the experimental values. The scissors operator cuts the LDA energy spectrum between the band gap and shifts the conduction band energies by a constant Δ to match the experimental band gap. The wavefunctions remain the same as the LDA wavefunctions. The corresponding Hamiltonian is given in Eq. (1.25) in the text.

"failure" of the scissors operators in the wide gap materials is not yet understood. To confirm the effects of the scissors operators, we need to take into account all the contributions which can lead to a 10% variation in the optical response. In particular, the local field, which gives about 5% correction to ε_∞ and about 20% correction to $\chi^{(2)}$, has to be considered.

1.6 Understanding local-field corrections using general response functions

In inhomogeneous systems, the polarization of the electrons in a spatially uniform external field induces another field – the "local field" – that varies in space. The polarization of the electrons depend on the total field – the macroscopic field plus the local field. The relative contribution to the macroscopic polarization from the local field is called the "local-field correction." In this section, we first provide the classical Clausius-Mossoti relation[43] for the local-field corrections to the linear and second-harmonic susceptibilities. Next, we describe the local-field corrections using the general response coefficients.[44] In particular, we show that the nonlinear-local-field correction to the second-harmonic susceptibility is directly related to the linear-local-field corrections. The microscopic descriptions of the local fields is deferred until Sec. 2.6.

From the classical electrostatics, the Clausius-Mossotti (or Lorentz-Lorenz) relation for the local-field corrections to the linear susceptibility $\chi^{(1)}$ can be derived for ionic and rare gas crystals.[43] In these systems where the electrons are tightly bound to the nucleus, the electric field on the electrons can be approximated by the electric field (E_I) at the position of the ions. For crystals with cubic symmetry, it can be shown from electrostatics that the electric field at the nucleus is[43]

$$E_I = E + 4\pi P/3, \tag{1.26}$$

where E is the spatially averaged electric field and P is the macroscopic polarization. Let $\chi_0^{(1)}$ be the linear susceptibility of individual ions:

$$P = \chi_0^{(1)} E_I. \tag{1.27}$$

From Eqs. (1.26) and (1.27), we obtain macroscopic linear susceptibility $\chi^{(1)}$

$$\chi^{(1)} = \frac{\chi_0^{(1)}}{1 - 4\pi\chi_0^{(1)}/3}, \tag{1.28}$$

where the factor $1/(1 - 4\pi\chi_0^{(1)}/3)$ comes from the local-field correction. This classical model informs us that the relative size of the variation of the electric field in the crystal is $4\pi\chi^{(1)}/3$, which, if we take $\chi^{(1)} = 0.2$, is of order unity. With similar analysis, the macroscopic second-harmonic susceptibility $\chi^{(2)}$ is related to the second-harmonic susceptibility of individual ions $\chi_0^{(2)}$ by[45,46]

$$\chi^{(2)} = \frac{\chi_0^{(2)}}{\left(1 - 4\pi\chi_0^{(1)}/3\right)^3}, \tag{1.29}$$

where the factor $1/(1 - 4\pi\chi_0^{(1)}/3)^3$ comes from the local-field corrections.

In covalent-bonded semiconductors, where the electrons are delocalized, the Clausius-Mossotti relation no longer applies. To obtain the local-field corrections in the covalent-bonded systems, we need to compute the induced microscopic density. With the induced microscopic densities $n^{(i)}(\vec{r})$ (the superscript denotes the order of the response in terms of the external field) the local fields $\phi^{(i)}(\vec{r})$ can be obtained from the Hartree and the exchange-correlation potentials rom the induced density

$$\phi^{(i)}(\vec{r}) = \int d^3r' \left[\frac{1}{|\vec{r} - \vec{r}\,'|} + \frac{v_{xc}[n_0(\vec{r}\,')]}{dn(\vec{r}\,')}\delta(\vec{r} - \vec{r}\,')\right] n^{(i)}(\vec{r}\,'), \quad , i = 1, 2, \cdots, \tag{1.30}$$

where n_0 is the ground state density. We point out that the size of the local field in covalent bonded crystals is on the same order as the average field. The local field in diamond from a tight-binding calculation is shown in Fig 1.5 as a sample case.[47]

For the second-harmonic susceptibility, there are two kinds of local-field corrections, those the linear and nonlinear (second-order) in the external field.[48] The

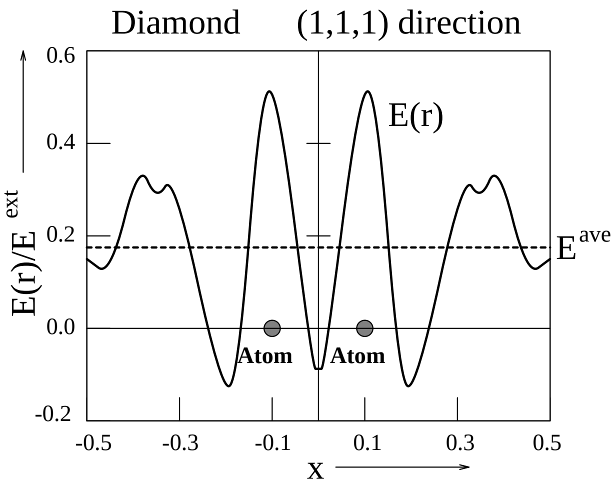

FIGURE 1.5 The local field in diamond (1,1,1) direction from a tight-binding calculation.[47] The solid curve gives the total electric field which is the sum of the local field and the average field (dashed line). The size of the local field is the same order as the average field. Note that in the some regions, the total electric field can be negative, i.e., in the opposite direction as the average field.

numerical procedure to calculate the local fields and the local-field corrections is postponed until Sec. 2.5. For the rest of this section, we try to understand the local fields using the general response functions.[44] In particular, we estimate the size of the nonlinear local-field correction which was argued to be negligible in a previous study.* The general response coefficients are denoted as $\chi^{(i)}$ with the superscript

* In Appendix B of Ref. 27, the argument for neglecting the nonlinear local fields used Maxwell's wave equation to estimate the size of the nonlinear local field $\vec{E}_s(\vec{r}, 2\omega)$. The wave equation was written as $-\nabla^2 \vec{E}_s(\vec{r}, 2\omega) - \frac{4\omega^2}{c^2} \vec{E}_s(\vec{r}, 2\omega) = \frac{4\pi}{c^2} \vec{P}(\vec{r}, 2\omega)$, where $\vec{P}(\vec{r}, 2\omega)$ is the nonlinear polarization at position $\vec{r}$. An implicit assumption of transverse electric field was made in the above equation to replace $\vec{\nabla} \times \vec{\nabla} \times \vec{E}_s(\vec{r}, 2\omega)$ by $-\nabla^2 \vec{E}_s(\vec{r}, 2\omega)$. This assumption is invalid because $\vec{E}_s(\vec{r}, 2\omega)$ arises from the Coulomb and exchange-correlation potential of the short-wave induced density and is longitudinal.

giving the order of the response, the leftmost subscript giving the induced quantity ($j \rightarrow$ current , $\rho \rightarrow$ density), and the second (and third when present) subscript from the left giving the driving potential ($j \rightarrow A$, $\rho \rightarrow \phi$). Figure 1.6 shows the relations between the induced currents/densities, the potentials, and the general response functions.

Linear susceptibility. We first use the general response functions to discuss the local-field corrections to the linear susceptibility. The driving sources are the spatially uniform electric field represented by the vector potential A and the linear local field $\phi^{(1)}$, where the spatial dependence is suppressed for simplicity. The linearly induced current $J^{(1)}$ is given by

$$J^{(1)} = \chi_{jj}^{(1)} A + \chi_{j\rho}^{(1)} \phi^{(1)}, \tag{1.31}$$

where the local-field correction comes from the second term $\chi_{j\rho}^{(1)} \phi^{(1)}$. To obtain the local field, we need to compute the linearly induced density

$$n^{(1)} = \chi_{\rho j}^{(1)} A + \chi_{\rho\rho}^{(1)} \phi^{(1)}. \tag{1.32}$$

For simplicity, we use V to represent both the hartree and exchange-correlation potentials and rewrite Eq. (1.30) by

$$\phi^{(i)} = V n^{(i)}, \qquad i = 1, 2, \cdots. \tag{1.33}$$

The density $n^{(1)}$ and the local field $\phi^{(1)}$ are coupled through Eqs. (1.32) and (1.33), which we can symbolically solve to get

$$\phi^{(1)} = \tilde{V} \chi_{\rho j}^{(1)} A, \tag{1.34}$$

where

$$\tilde{V} \equiv \frac{V}{1 - V\chi_{\rho\rho}^{(1)}}, \tag{1.35}$$

is the screened potential. Substituting $\phi^{(1)}$ in Eq. (1.31) by Eq. (1.34) yields

$$J^{(1)} = \chi_{jj}^{(1)} A + \chi_{j\rho}^{(1)} \tilde{V} \chi_{\rho j}^{(1)} A, \tag{1.36}$$

where the second term on the right hand side gives the local-field correction to the induced macroscopic current.

Second-harmonic susceptibility. In analogy with the linear case, the expression for the second-order nonlinearly induced current is given by

$$J^{(2)} = \chi_{jjj}^{(2)} AA + \left(\chi_{jj\rho}^{(2)} + \chi_{j\rho j}^{(2)}\right) A\phi^{(1)} + \chi_{j\rho\rho}^{(2)} \phi^{(1)}\phi^{(1)} + \underline{\chi_{j\rho}^{(2)} \phi^{(2)}}, \tag{1.37}$$

and the nonlinearly induced density is given by

$$n^{(2)} = \chi_{\rho jj}^{(2)} AA + \left(\chi_{\rho j\rho}^{(2)} + \chi_{\rho\rho j}^{(2)}\right) A\phi^{(1)} + \chi_{\rho\rho\rho}^{(2)} \phi^{(1)}\phi^{(1)} + \chi_{\rho\rho}^{(2)} \phi^{(2)}. \tag{1.38}$$

We underline the contribution to $J^{(2)}$ from the nonlinear local field $\phi^{(2)}$, which was argued to be negligible in a previous study (see footnote on p.24). However, as we will show below, this contribution from the nonlinear-local-field correction is as important as the contributions from the linear local field. To see this, we may express $\phi^{(2)}$ in terms of A and $\phi^{(1)}$. The nonlinearly induced density $n^{(2)}$ and the nonlinear local field $\phi^{(2)}$ are coupled through Eqs. (1.38) and (1.33), which can be symbolically solved to give

$$\phi^{(2)} = \tilde{V} \left[\chi_{\rho jj}^{(2)} AA + \left(\chi_{\rho j\rho}^{(2)} + \chi_{\rho\rho j}^{(2)}\right) A\phi^{(1)} + \chi_{\rho\rho\rho}^{(2)} \phi^{(1)}\phi^{(1)}\right]. \tag{1.39}$$

Substituting $\phi^{(2)}$ in Eq. (1.37) by Eq. (1.39) and observing that $\chi_{j\rho}^{(1)} A = \chi_{\rho j}^{(1)} A =$

$\phi^{(1)}$ from Eq. (1.34) ($\chi^{(1)}_{j\rho} = \chi^{(1)}_{\rho j}$ will be proved later) yields

$$
\begin{aligned}
J^{(2)} =& \chi^{(2)}_{jjj} AA + \chi^{(2)}_{jj\rho} A\phi^{(1)} + \chi^{(2)}_{j\rho j}\phi^{(1)} A + \chi^{(2)}_{j\rho\rho}\phi^{(1)}\phi^{(1)} \\
&+ \tilde{V}\left[\underline{\chi^{(2)}_{\rho jj}} AA + \left(\underline{\chi^{(2)}_{\rho j\rho}} + \underline{\chi^{(2)}_{\rho\rho j}}\right) A\phi^{(1)} + \underline{\chi^{(2)}_{\rho\rho\rho}}\phi^{(1)}\phi^{(1)}\right] \\
=& \chi^{(2)}_{jjj} AA && \text{(long-wave)} \\
&+ \left[\chi^{(2)}_{jj\rho} + \chi^{(2)}_{j\rho j} + \underline{\chi^{(2)}_{\rho jj}}\right] A\phi^{(1)} && \text{(first-order)} \\
&+ \left[\chi^{(2)}_{j\rho\rho} + \underline{\chi^{(2)}_{\rho j\rho}} + \underline{\chi^{(2)}_{\rho\rho j}}\right]\phi^{(1)}\phi^{(1)} && \text{(second-order)} \\
&+ \underline{\chi^{(2)}_{\rho\rho\rho}}\phi^{(1)}\phi^{(1)}\phi^{(1)}/A, && \text{(third-order)}
\end{aligned}
\tag{1.40}
$$

where the terms arising from $\chi^{(1)}\phi^{(2)}$ in Eq. (1.37) are underlined. Thus, we have transformed the nonlinear-local-field correction to the linear-local-field corrections.

To further simplify the relation between the nonlinear- and linear-local-field corrections, we use an intrinsic symmetry between the response functions,

$$
\begin{aligned}
\chi^{(1)}_{j\rho} &= \chi^{(1)}_{\rho j}, \\
\chi^{(2)}_{jj\rho} &= \chi^{(2)}_{j\rho j} = \chi^{(2)}_{\rho jj}, \\
\chi^{(2)}_{j\rho\rho} &= \chi^{(2)}_{\rho\rho j} = \chi^{(2)}_{\rho j\rho},
\end{aligned}
\tag{1.41}
$$

which is proved later in this section. Combining Eqs. (1.40) and (1.41), we see that to first-order correction (i.e., terms including the first power of the linear local field), the term from the nonlinear local field (underlined) is equal to half the terms from the linear local field. Similarly, to second-order correction (i.e., terms including the second power of the linear local field), the terms from the nonlinear local field (underlined) are equal to twice the term from the linear local field. In addition, there is third-order correction from the nonlinear local field. As we will show in Sec. 3.4, the dominant correction comes from the first-order terms. Thus the nonlinear local field adds 50% increase to the linear-local-field corrections to

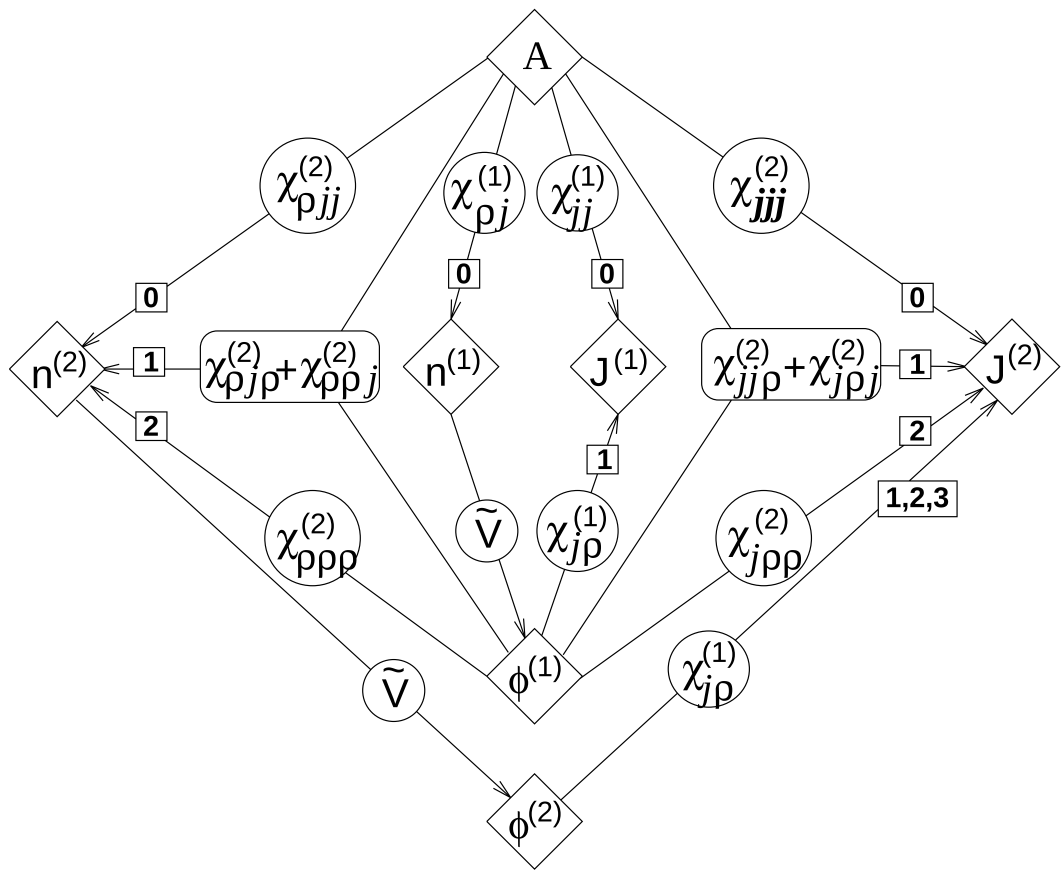

FIGURE 1.6 Flow chart showing the relations between the various potentials, induced currents/densities, and general response functions. The quantities written in the diamonds include the vector potential A, the local fields $\phi^{(i)}$, and the induced currents $J^{(i)}$, and induced densities $n^{(i)}$, with $i = 1, 2$ denoting the first- and second-order response. Linear and nonlinear ($i = 1, 2$ in the superscripts) general response functions χ are written in circles (and boxy circles) associated with each line relating a potential and a current/density. The first subscript from the left denotes the current/density and the second (and third when present) subscript from the left denotes the potential. The numbers written in the boxes on the arrows give the different orders in the local-field corrections, where 0 stands for the long-wave contribution. The screened potential $\tilde{V} = V/(1 - V\chi^{(1)}_{\rho\rho})$ [cf. Eq. (1.35)] gives the self-consistency loop between the induced densities and the corresponding local fields. The relations shown in this figure are described by Eqs. (1.31), (1.32), (1.37) and (1.38) in the text.

the second-harmonic susceptibility. The microscopic theory given in Sec. 2.10 gives identical conclusions.

To understand the intrinsic symmetry of the general response coefficients, consider the change of the Kohn-Sham total energy functional $E[n, J]$ as a function of the induced current $J = J^{(1)} + J^{(2)}$ and induced density $n = n^{(1)} + n^{(2)}$ in the presence of an external field represented by the vector potential A. We extend LDA to include a current dependence in the kinetic energy functional $T[n, J]$. We use $T''_{j\rho} \equiv \partial^2 T/\partial J \partial n$ for the partial derivative and use K and K' to denote the hartree-exchange-correlation kernel and its first derivative. The change in the total energy functional is, to the third-order in the induced current J and the induced density n,

$$\begin{aligned} E[n, J] =& \frac{1}{2!} T''_{jj} J^2 + T''_{j\rho} J n + \frac{1}{2!} T''_{\rho\rho} n^2 + K n^2 + A J \\ & + \frac{1}{3!} T'''_{jjj} J^3 + \frac{1}{2} T'''_{jj\rho} J^2 n + \frac{1}{2} T'''_{j\rho\rho} J n^2 + \frac{1}{3!} T'''_{\rho\rho\rho} n^3 + \frac{1}{3} K' n^3. \end{aligned} \tag{1.42}$$

All derivatives in the above equation are evaluated at the ground-state density. The induced current and density correspond to the values that minimize $E[n, J]$:

$$\begin{aligned} 0 &= \frac{\delta E[n, J]}{\delta J} = T''_{jj} J + T''_{j\rho} n + A + \frac{1}{2} T'''_{jjj} J^2 + T'''_{jj\rho} J n + \frac{1}{2} T'''_{j\rho\rho} n^2, \\ 0 &= \frac{\delta E[n, J]}{\delta n} = T''_{j\rho} J + T''_{\rho\rho} n + 2 K n + \frac{1}{2} T'''_{jj\rho} J^2 + T'''_{j\rho\rho} J n + \frac{1}{2} T'''_{\rho\rho\rho} n^2 + K' n^2. \end{aligned} \tag{1.43}$$

To set up a connection between the general response functions and the kinetic and hartree-exchange-correlation energy functionals, we seek solutions to Eq. (1.43) in the same from as Eqs. (1.31), (1.32), (1.37) and (1.38). The first- and second-order local fields acting on the non-interacting Kohn-Sham electrons are

$$\begin{aligned} \phi^{(1)} &= 2 K n^{(1)}, \\ \phi^{(2)} &= 2 K n^{(2)} + K' \left(n^{(1)}\right)^2. \end{aligned} \tag{1.44}$$

From Eqs. (1.43) and (1.44), we can obtain the equations for the linearly induced current and density ($T''_{\rho j} = T''_{j\rho}$)

$$
\begin{aligned}
J^{(1)} &= \frac{1}{D}\left(T''_{jj}A - T_{j\rho}\phi^{(1)}\right), \\
n^{(1)} &= \frac{1}{D}\left(T''_{\rho j}A - T_{\rho\rho}\phi^{(1)}\right),
\end{aligned}
\tag{1.45}
$$

where

$$
D \equiv T''_{jj}T''_{\rho\rho} - T''_{j\rho}T''_{\rho j}. \tag{1.46}
$$

Comparing Eq. (1.45) with the definitions of the general response functions given in Eqs. (1.31) and (1.32) yields

$$
\chi^{(1)}_{j\rho} = \frac{T''_{j\rho}}{D} = \frac{T''_{\rho j}}{D} = \chi^{(1)}_{\rho j}, \tag{1.47}
$$

which is the intrinsic symmetry of the first-order general response functions. With some algebra, the intrinsic symmetry of second-order general response functions can be proved analogously,

$$
\begin{aligned}
&\chi^{(2)}_{jj\rho} = \chi^{(2)}_{j\rho j} = \chi^{(2)}_{\rho jj} \\
=&\frac{1}{2D^3}\left[-T'''_{jjj}T''_{j\rho}(T''_{\rho\rho})^2 + 2T'''_{jj\rho}T''_{\rho\rho}(T''_{j\rho})^2 + T'''_{jj\rho}T''_{jj}(T''_{\rho\rho})^2 + T'''_{\rho\rho\rho}T''_{jj}(T''_{j\rho})^2\right], \\
&\chi^{(2)}_{j\rho\rho} = \chi^{(2)}_{\rho\rho j} = \chi^{(2)}_{\rho j\rho} \\
=&\frac{1}{2D^3}\left[T'''_{jjj}T''_{\rho\rho}(T''_{j\rho})^2 - 2T'''_{jj\rho}T''_{jj}T''_{\rho\rho}T''_{\rho\rho} - T'''_{jj\rho}(T''_{j\rho})^3 - T'''_{\rho\rho\rho}T''_{j\rho}(T''_{jj})^2\right].
\end{aligned}
\tag{1.48}
$$

This completes the proof of the intrinsic symmetry of the general response functions.

By using the general response functions, we have algebraically transformed the nonlinear-local-field corrections to the second-harmonic susceptibility to the linear-local-field corrections. With the intrinsic symmetry between the response functions, the relation between the nonlinear- and linear-local-field corrections simplifies to a factor of 1/2 for the first-order correction and a factor of 2 for the second-order correction. As will be shown in Sec. 2.6, these relations can also be proved from the microscopic expressions. The general response functions are useful in understanding the relations between the different local fields, however, they do not offer any information on the size of the local-field corrections. We need to carry out numerical calculations to compute the local-field corrections from the microscopic expressions. Results will be shown in Sec. 3.4.

1.7 Organization for the rest of the thesis

In Chapter 2, we give a detailed derivation for the microscopic expressions for the linear and second-harmonic susceptibilities. The self-energy correction is included in the form of the scissors operators. For the second-harmonic susceptibility, there are two kinds of local-field corrections, those linear and nonlinear in the external field. Using the microscopic expressions, we demonstrate that the previously neglected nonlinear-local-field corrections are as important as the linear-local-field corrections. We establish a direct relation between the nonlinear- and linear-local-field corrections, which, in the static limit, reduces to a factor of 1/2 for the first-order correction and a factor of 2 for the second-order correction.

In Chapter 3, we carry out systematic studies of the linear and second-harmonic susceptibilities in a series of wide gap semiconductors including diamond, four polytypes of SiC (2H, 4H, 6H, and 3C), BN, AlN, and GaN. The pseudopotentials we use for these studies are carefully constructed and tested to give reliable and consistent results as other methods (e.g., the linear-augmented-plane-wave method) and other pseudopotentials. The four polytypes of SiC, which have identical chemical composition, differ in the stacking sequence of the atomic layers. Through the study of the linear and second-harmonic susceptibilities, we may gain insight into the structural dependence of the optical properties. For the III-V nitrides, both the zincblende and wurtzite structures are studied. We find that, although the linear susceptibilities are not sensitive the structural changes in these compounds, the second-harmonic susceptibilities can differ by as large as a factor of ten between the zincblende and wurtzite structures. This structural dependence cannot be explained by any of the semi-classical models for the second-harmonic susceptibility.[6,9] Through our current studies on diamond, SiC, BN, AlN, GaN, and the previously studied urea,[30] we find that the scissors-operator approach to calculate optical response coefficients, which is successful in small to medium gap materials,[32] gives values 5-20% smaller than experiment for these wide gap semiconductors.[49,50,30] The nonlinear-local-field corrections are calculated for most of the materials previously studied. The numerical results show that, for the second-harmonic susceptibility, the local-field corrections form a hierarchy structure among the three orders of corrections given in Eq. (1.40). The leading correction comes from the first-order terms, 2/3 of which arises from the linear local field and 1/3 from the nonlinear local field.

In Chapter 4, we summarize the results on the linear and second-harmonic susceptibilities. We give some possible reasons for the failure of the scissors operators in the wide gap materials.

In Appendix A, we provide a general scheme to generate the special $\vec{k}$-points for Brillouin zone integrations. In Appendix B, we describe a method to compute the spin-orbit coupling using pseudopotentials. In Appendix C, we give a detailed proof on the equivalence of the newly proposed "$2n+1$" method and the traditional sum-over-states method for the second-harmonic susceptibility.

REFERENCES

1. Y. R. Shen, *The principles of nonlinear optics,* (New York, Wiley 1984).
2. P. A. Franken, A. e. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Lett. **7**, 118 (1961).
3. *CRC handbook of laser science and technology,* edited by M. J. Weber, (Boca Raton, Fla) 1982.
4. S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, Jap. J. Appl. Phys. Part 2-lett. v35, L74 (1996); v35, L217 (1996); Appl. Phys. Lett. **68**, 2105 (1996); **68**, 3269 (1996).
5. For a list of other nonlinear crystals and the conversion efficiencies, see V. G. Dmitriev, G. G. Gurzadyan, and D. N. Nidogosyan, *Handbook of nonlinear optical crystals,* (Springer-Verlag) 1991.
6. B. F. Levine, Phys. Rev. B **7**, 2600 (1973).
7. D. A. Kleinman, Phys. Rev. B **2**, 3139 (1970).
8. F. N. H. Robinson, Phys. Lett. **26A**, 435 (1968).
9. D. A. Kleinman, Phys. Rev. **126**, 1977 (1962).
10. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
11. W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).
12. E. P. Wigner, Trans. Faraday Soc. **34**, 678 (1938).
13. D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
14. M. P. Teter, private communication.
15. For a review of the density functional method and its applications, see R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
16. J. P. Perdew and K. Burke, Int. J. Quant. Chem. **57**, 309 (1996).
17. A. Dal Corso, S. Baroni, and R. Resta, Phys. Rev. B **49**, 5323 (1994).
18. W. Kohn and P. Vashishta, *Theory of the inhomogeneous electron gas,* edited by S. Lundqvist and N. H. March, (Plenum Press, New York, 1983). p92.
19. P. N. Butcher and T. P. McLean, Proc. Phys. Soc. (London) **81**, 219 (1963); **83**, 579 (1964).
20. D. E. Aspnes, Phys. Rev. B **6**, 4648 (1972).
21. C. Y. Fong and Y. R. Shen, Phys. Rev. B **12**, 2325 (1975).
22. D. J. Moss, E. Ghahramani, J. E. Sipe, and H. M. van Driel, Phys. Rev. B **41**, 1542 (1990).

23. E. Ghahramani, D. J. Moss, and J. E. Sipe, Phys. Rev. B **43**, 8990 (1991).

24. H. Ma, S. T. Chui, R. V. Kasowski, and W. Y. Hsu, Opt. comm, **85**, 437 (1991).

25. M-Z. Huang and W. Y. Ching, Phys. Rev. B **45**, 8738 (1992).

26. M-Z. Huang and W. Y. Ching, Phys. Rev. B **47**, 9464 (1993).

27. Z. H. Levine and D. C. Allan, Phys. Rev. B **44**, 12781 (1991); **48**, 14768 (E) (1993).

28. Z. H. Levine, and D. C. Allan, Phys. Rev. Lett. **16**, 1719 (1989); Phys. Rev. B **43**, 4187 (1991).

29. H. Zhong, Z. H. Levine, D. C. Allan, and J. W. Wilkins, Phys. Rev. B **48**, 1384 (1993).

30. Z. H. Levine and D. C. Allan, Phys. Rev. B **48**, 7783 (1993).

31. A. Dal Corso and F. Mauri, Phys. Rev. B **50**, 5756 (1994); A. Dal Corso, F. Mauri, and A. Rubio, *ibid.* **53**, 15638 (1996).

32. For an overall review of the previous studies, see Z. H. Levine, Int. J. Quantum Chem. **S28**, 411 (1994).

33. J. J. Sham and M. Schlüter, Phys. Rev. B **32**, 3883 (1985).

34. L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1969).

35. M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986). [C, Si, Ge, and LiCl]

36. S. L. Adler, Phys. Rev. **126**, 413 (1962); N. Wiser, *ibid.* **129**, 62 (1962).

37. R. W. Godby, M. Schlüter, and L. J. Sham Phys. Rev. B **37**, 10159 (1988). [Si, GaAs, AlAs, and C]

38. M. P. Surh, S. G. Louie, and M. L. Cohen, Phys. Rev. B **43**, 9126 (1991). [BN, BP, and BAs]

39. A. Rubio, J. L. Corkill, M. L. Cohen, E. L. Shirley, and S. G. Louie, Phys. Rev. B **48**, 11810 (1993). [AlN and GaN]

40. M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **48**, 17791 (1993). [C, Si, Ge, GaAs, and 3C-SiC]

41. B. Wenzien, P. Käckell, and F. Bechstdet, Phys. Rev. B **52**, 10897 (1995). [2H-, 4H, 6H, and 3C-SiC]

42. S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).

43. N. W. Ashcroft and N. D. Mermin, *Solid state physics,* (New York, Holt, Rinehart and Winston, 1976), p539-542.

44. L. Jönsson, Z. H. Levine, and J. W. Wilkins, Phys. Rev. Lett. **76**, 1372 (1996).

45. N. Bloembergen, *Nonlinear optics,* (Benjamin, New York, 1965.)

46. V. Mizrahi and J. E. Sipe, Phys. Rev. B **34**, 3700 (1986).

47. L. Oliveira, Phys. Stat. Sol. (b) **107**, 255 (1981).

48. J. Chen, L. Jönsson, J. W. Wilkins, and Z. H. Levine, submitted to Phys. Rev. B.

49. J. Chen, Z. H. Levine, and J. W. Wilkins, Phys. Rev. B **50**, 11514 (1994).

50. J. Chen, Z. H. Levine, and J. W. Wilkins, Appl. Phys. Lett. **66**, 1129 (1995).