

Electronic band structure from first-principles Green's function approach: theory and implementations

Hong JIANG (✉)

Electronic band structure is one of the most important intrinsic properties of a material, and is in particular crucial in electronic, photo-electronic and photo-catalytic applications. Kohn-Sham Density-functional theory (KS-DFT) within currently available local or semi-local approximations to the exchange-correlation energy functional is problematic for the description of electronic band structure. Many-body perturbation theory based on Green's function (GF) provides a rigorous framework to describe excited-state properties of materials. The central ingredient of the GF-based many-body perturbation theory is the exchange-correlation self-energy, which accounts for all non-classical electron-electron interaction effects beyond the Hartree theory, and formally can be obtained by solving a set of complicated integro-differential equations, named Hedin's equations. The *GW* approximation, in which the self-energy is simply a product of Green's function and the screened Coulomb interaction (*W*), is currently the most accurate first-principles approach to describe electronic band structure properties of extended systems. Compared to KS-DFT, the computational efforts required for *GW* calculations are much larger. Various numerical techniques or approximations have been developed to apply *GW* for realistic systems. In this paper, we give an overview of the theory of first-principles Green's function approach in the *GW* approximation and review the state of the art for the implementation of *GW* in different representations and with different treatment of the frequency dependence. It is hoped that further methodological developments will be inspired by this work so that the approach can be applied to more complicated and scientifically more interesting systems.

Received November 6, 2011; accepted November 23, 2011
Beijing National Laboratory for Molecular Sciences, State Key Laboratory of Rare Earth Material Chemistry and Application, Institute of Theoretical and Computational Chemistry, College of Chemistry, Peking University, 100871 Beijing, China
E-mail: h.jiang@pku.edu.cn

Keywords electronic band structure, many-body perturbation theory, *GW* approximation

1 Introduction

Electronic band structure is one of the most important intrinsic properties of a material that has far-reaching effects on electronic, optical, photo-electronic and photo-catalytic applications. Experimentally electronic properties of a material of interest are often measured in terms of its response to some external perturbation, using either electron or photon as the probe. Among the most widely used spectroscopic techniques are photo-emission, inverse photo-emission and optical absorption spectroscopies [1–3]. In the photo-emission spectroscopy (PES) [1], photons of given energy impinge on the sample and liberate electrons from their occupied levels, either deep core states or valence states; by monitoring the kinetic energy of photoelectrons, the information of electronic occupied states can be probed. The inverse photo-emission spectroscopy (IPS) is essentially a time-reversed process of PES: free electrons of given energy are injected into the sample and fill originally unoccupied states; the redundant energy is released as radiation, whose intensity as a function of energy reveals the information of unoccupied (or conduction band) states. In essence, PES and IPS measure single-particle excitations of the system, with the total electron number N changing to $N - 1$ or $N + 1$, which are often termed as quasiparticle (QP) excitations. The optical absorption spectroscopy (OAS), on the other hand, measures neutral excitations, in which electrons are excited from the ground state to excited states, corresponding to creating a hole and an electron simultaneously in a QP picture. If we neglect the interaction between photo-excited electrons and holes, OAS can be regarded as the combination of PES and IPS in a single process. In this paper, we are mainly concerned with the theoretical description of QP excitations from a first-principles perspective.

Kohn-Sham (KS) density functional theory (DFT) [4,5] within the local density or generalized gradient approximation (LDA/GGA) to the exchange-correlation (xc) energy functional has become “the standard model” for first-principles electronic structure calculations of extended systems [6]. By mapping the many-electron interacting system to a fictitious non-interacting (Kohn-Sham) system, which has the same ground state electron density as the interacting one, the original highly complicated many-body problem is transformed to solving a single-particle equation that formally resembles a mean-field theory, but is exact in principles. Within LDA/GGA, KS-DFT can provide accurate descriptions of energetic and structural properties for many materials with feasible computational efforts. As a byproduct, KS-DFT

also gives a set of single-particle energies and wave functions, which are, rigorously speaking, auxiliary quantities introduced only to calculate electron density, and therefore do not have any physical meanings. In practice, however, they are often used to interpret electronic quasi-particle band structure of materials as probed by PES/IPS [1]. This practice, however, has to be exercised with caution. Even for many simple *sp* semiconductors, KS-DFT within LDA/GGA gives band gaps that are systematically underestimated when compared to experiment, and can predict metallic ground state for small-gap semiconductors, e.g., Ge, InN and so on [7].

A completely different theoretical framework for electronic band structure is provided by many-body perturbation theory (MBPT), formulated in terms of one-body Green's function [8,9], in the *GW* approximation [10,11]. In this approach, the exchange-correlation self-energy Σ , the central quantity in GF-based MBPT, is obtained as a simple product of single-particle Green's function (*G*) and the screened Coulomb interaction (*W*), hence termed *GW*. In practice, quasi-particle energies in the *GW* approximation are often calculated as a first-order correction to eigenenergies of some reference single-particle Hamiltonian H_0 , and both *G* and *W* are calculated using eigen-energies and eigen-functions of H_0 , hence called G_0W_0 . Since the middle of 1980s [12,13], the G_0W_0 approach based on the LDA or GGA H_0 has become the method of choice for quasi-particle band structures of various semiconductors [11]. From a practical point of view, the computational demand of *GW* calculations is much heavier than KS-DFT within LDA/GGA. As a result, various physical or numerical approximations have been developed to facilitate the applications of *GW* to physically interesting systems. Recent years have seen intensive developments in the methodology of *GW* with the goal of either going beyond LDA-based G_0W_0 [14–16], or significantly extending the capability of current implementations so that more complex systems can be reached [17–21]. The main goal of this paper is to present an overview of many-body perturbation theory for quasi-particle excitations based on the Green's function, and review various numerical techniques developed for the implementation of the *GW* approach. The topic covered in this paper partly overlap with that in Ref. [22], but with different focus. Considering the recent increasing interest to apply the *GW* for large systems that are currently out of reach of current implementations, it is hoped that this work will inspire further methodological developments so that routine applications of *GW* and related first-principles methods to more complicated and scientifically more interesting systems become feasible.

The paper is organized as follows. In the next section we present the general theoretical framework in which the *GW* approximation is derived. In Section 3, main numerical techniques used in the implementation of the *GW* method are

discussed including the basis used to represent the *GW* equations and the treatment of the frequency dependence by different approaches. Section 4 concludes the paper with a few general remarks.

2 Many-body perturbation theory and the *GW* approximation

In this section we present the fundamentals of first principles many-body perturbation theory based on one-body Green's function. More complete formalisms can be found in Refs. [9,11,23,24].

2.1 Green's function

The Hamiltonian of a many-electron interacting system, which is determined by the external potential $V_{\text{ext}}(\mathbf{x})$ and the number of electrons (*N*), reads in the second-quantization representation [8]

$$\hat{H} = \int d\mathbf{x} \hat{\psi}^\dagger(\mathbf{x}t) h_0(\mathbf{x}) \hat{\psi}(\mathbf{x}t) + \frac{1}{2} \int d\mathbf{x} d\mathbf{x}' v(\mathbf{x}, \mathbf{x}') \hat{\psi}^\dagger(\mathbf{x}t) \hat{\psi}^\dagger(\mathbf{x}'t) \hat{\psi}(\mathbf{x}'t) \hat{\psi}(\mathbf{x}t) \quad (1)$$

where we use $\mathbf{x} \equiv \{\mathbf{r}, \sigma\}$ to denote both spatial and spin coordinates. $\hat{\psi}(\mathbf{x}t)$ and $\hat{\psi}^\dagger(\mathbf{x}t)$ are the annihilation and creation field operators in the Heisenberg picture, respectively. $h_0 = \frac{1}{2} \nabla^2 + V_{\text{ext}}(\mathbf{x})$ is the one-body part of the Hamiltonian, and $v(\mathbf{x}, \mathbf{x}') \equiv \frac{1}{|\mathbf{r} - \mathbf{r}'|} \delta_{\sigma, \sigma'}$ is the bare Coulomb interaction, which is spin-independent. Atomic units are used through the paper.

One-body Green's function $G(\mathbf{x}t, \mathbf{x}'t')$, also called the propagator, is then defined as [8]

$$G(\mathbf{x}t; \mathbf{x}'t') = -i \langle N | \mathbf{T} \left[\hat{\psi}(\mathbf{x}t) \hat{\psi}^\dagger(\mathbf{x}'t') \right] | N \rangle \quad (2)$$

where \mathbf{T} is the time-ordering operator, and $|N\rangle$ denotes the ground state of the *N*-electron interacting system. Based on the Heisenberg equation of motion (EOM) for the field operator $\hat{\psi}(\mathbf{x}t)$

$$i \frac{\delta \hat{\psi}(\mathbf{x}t)}{\delta t} = [\hat{\psi}(\mathbf{x}t), \hat{H}] \quad (3)$$

we can obtain the equation of motion for the Green's function

$$\begin{aligned} & \left[i \frac{\delta}{\delta t} - h_0(\mathbf{x}) \right] G(\mathbf{x}t, \mathbf{x}'t') - i \int d\mathbf{x}'' v(\mathbf{x}, \mathbf{x}'') \\ & \times \langle N | \mathbf{T} \left[\hat{\psi}^\dagger(\mathbf{x}''t) \hat{\psi}(\mathbf{x}''t) \hat{\psi}(\mathbf{x}t) \hat{\psi}^\dagger(\mathbf{x}'t) \right] | N \rangle \\ & = \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \end{aligned} \quad (4)$$

The last term in the left side of the equation above is a special case of the two-body Green's function, generally defined as

$$G_2(\mathbf{x}_1 t_1, \mathbf{x}_2 t_2, \mathbf{x}_3 t_3, \mathbf{x}_4 t_4) = i^2 \langle N | \mathbf{T} \left[\hat{\psi}(\mathbf{x}_1 t_1) \hat{\psi}(\mathbf{x}_3 t_3) \hat{\psi}^\dagger(\mathbf{x}_4 t_4) \hat{\psi}^\dagger(\mathbf{x}_2 t_2) \right] | N \rangle \quad (5)$$

Equation (4) can therefore be written as

$$\begin{aligned} & \left[i \frac{\delta}{\delta t} - h_0(\mathbf{x}) \right] G(\mathbf{x}t, \mathbf{x}'t') + i \int d\mathbf{x}'' v(\mathbf{x}, \mathbf{x}'') \\ & \times G_2(\mathbf{x}t, \mathbf{x}'t', \mathbf{x}''t, \mathbf{x}''t^+) \\ & = \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \end{aligned} \quad (6)$$

with $t^+ \equiv t + \eta$, where η is an infinitesimal positive number. The equation of motion for G_2 , which depends on the three-body Green's function, can be derived in a similar way. This process can continue, leading to a set of hierarchical equations that are formally exact but useless in practice. To obtain practically feasible approximations, it is necessary to cut off the series or reformulate the equations in a form that is more accessible by approximations.

Equation (6) can be used as the starting point for simple approximations [23] such as the Hartree approximation

$$G_2(\mathbf{x}t, \mathbf{x}'t', \mathbf{x}''t, \mathbf{x}''t^+) \simeq G(\mathbf{x}t, \mathbf{x}'t') G(\mathbf{x}''t, \mathbf{x}''t^+) \quad (7)$$

or the Hartree-Fock (HF) approximation

$$\begin{aligned} & G_2(\mathbf{x}t, \mathbf{x}'t', \mathbf{x}''t, \mathbf{x}''t^+) \\ & \simeq G(\mathbf{x}t, \mathbf{x}'t') G(\mathbf{x}''t, \mathbf{x}''t^+) + G(\mathbf{x}t, \mathbf{x}''t^+) G(\mathbf{x}''t, \mathbf{x}'t') \end{aligned} \quad (8)$$

It is, however, difficult to incorporate high order approximations directly based on Eq. (6). Formally it is much more convenient to reformulate Eq. (6) in terms of the exchange-correlation self-energy.

2.2 The self-energy and quasi-particle equation

The exchange-correlation self-energy Σ (or simply called self-energy) can be formally defined as

$$\begin{aligned} & i \int d\mathbf{x}'' v(\mathbf{x}, \mathbf{x}'') G_2(\mathbf{x}t, \mathbf{x}'t', \mathbf{x}''t, \mathbf{x}''t^+) \\ & \equiv -V_H(\mathbf{x}) G(\mathbf{x}t, \mathbf{x}'t') \\ & - \int d\mathbf{x}'' dt'' \Sigma(\mathbf{x}t, \mathbf{x}''t'') G(\mathbf{x}''t'', \mathbf{x}'t') \end{aligned} \quad (9)$$

where the first term on the right side contains the contribution of classical Coulomb repulsion (the Hartree potential), and all complexities related to the two-body and higher order Green's functions are now wrapped up in the self-energy operator Σ . The equation of motion for the Green's function now reads

$$\begin{aligned} & \left[i \frac{\delta}{\delta t} - h_0(\mathbf{x}) - V_H(\mathbf{x}) \right] G(\mathbf{x}t, \mathbf{x}'t') \\ & - \int d\mathbf{x}'' dt'' \Sigma(\mathbf{x}t, \mathbf{x}''t'') G(\mathbf{x}''t'', \mathbf{x}'t') \\ & = \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \end{aligned} \quad (10)$$

For a time-independent system, both $G(\mathbf{x}t, \mathbf{x}'t')$ and $\Sigma(\mathbf{x}t, \mathbf{x}'t')$ depend on only the time difference $\tau = t - t'$. Theoretically it is more convenient to work in the frequency domain, obtained by taking a Fourier transform with respect to τ , which leads to

$$\begin{aligned} & [\omega - h_0(\mathbf{x}) - V_H(\mathbf{x})] G(\mathbf{x}, \mathbf{x}'; \omega) \\ & - \int d\mathbf{x}'' \Sigma(\mathbf{x}, \mathbf{x}''; \omega) G(\mathbf{x}'', \mathbf{x}'; \omega) \\ & = \delta(\mathbf{x} - \mathbf{x}') \end{aligned} \quad (11)$$

By introducing the non-interacting Green's function G_0 as the solution of

$$[\omega - h_0(\mathbf{x}) + V_H(\mathbf{x})] G_0(\mathbf{x}, \mathbf{x}'; \omega) = \delta(\mathbf{x} - \mathbf{x}') \quad (12)$$

Equation (11) can be re-written as

$$\begin{aligned} G(\mathbf{x}, \mathbf{x}'; \omega) & = G_0(\mathbf{x}, \mathbf{x}'; \omega) + \int d\mathbf{x}_1 d\mathbf{x}_2 G_0(\mathbf{x}, \mathbf{x}_1; \omega) \\ & \times \Sigma(\mathbf{x}_1, \mathbf{x}_2; \omega) G(\mathbf{x}_2, \mathbf{x}'; \omega), \end{aligned} \quad (13)$$

which is the Dyson's equation for the Green's function G .

The (one-body) Green's function contains a lot of important information related to the ground and excited states of the system [8]. The poles of $G(\mathbf{x}, \mathbf{x}'; \omega)$ in the complex frequency domain give single-particle excitation properties ($N \rightarrow N \pm 1$) as probed by PES/IPS. In addition, the ground state total energy can also be obtained from $G(\mathbf{x}, \mathbf{x}'; \omega)$ [8]. For the description of the single-particle excitation, it is more convenient to work directly with the so-called quasi-particle equation, whose solutions give QP energies and wave functions. Here we present a schematic derivation of the QP equation [11]. According to the Green's function theory of differential equations [25], $G(\mathbf{x}, \mathbf{x}'; \omega)$ as the solution of Eq. (11) can be formally written as

$$G(\mathbf{x}, \mathbf{x}'; \omega) = \sum_n \frac{\Psi_n(\mathbf{x}; \omega) \Psi_n^*(\mathbf{x}'; \omega)}{\omega - \varepsilon_n(\omega)} \quad (14)$$

where $\Psi_n(\mathbf{x}; \omega)$ and $\varepsilon_n(\omega)$ are the eigen-solutions of the following equation

$$\begin{aligned} & [h_0(\mathbf{x}) + V_H(\mathbf{x})] \Psi_n(\mathbf{x}; \omega) \\ & + \int d\mathbf{x}' \Sigma(\mathbf{x}, \mathbf{x}'; \omega) \Psi_n(\mathbf{x}'; \omega) \\ & = E_n(\omega) \Psi_n(\mathbf{x}; \omega) \end{aligned} \quad (15)$$

By definition, ω is real, but formally the domain of ω can be extended to the whole complex plane by analytic continuation, so that the poles of $G(\mathbf{x}, \mathbf{x}', \omega)$, with $\omega = \varepsilon_n(\omega)$, correspond to quasi-particle energies. Since QP energies and wave functions are the quantities of main interest, Eq. (15) can be rewritten in the following form, often denoted as the QP equation,

$$\begin{aligned} [h_0(\mathbf{x}) + V_H(\mathbf{x})]\Psi_n(\mathbf{x}) + \int d\mathbf{x}' \Sigma(\mathbf{x}, \mathbf{x}'; \varepsilon_n)\Psi_n(\mathbf{x}') \\ = \varepsilon_n \Psi_n(\mathbf{x}) \end{aligned} \quad (16)$$

We note that the term “quasi-particle” here is used in a generalized sense.

2.3 Hedin's equations and the GW approximation

The self-energy operator is itself a highly complicated quantity, and requires further approximations in practice. There are mainly two approaches to approximate Σ in a computationally accessible manner: 1) the Feynman's diagrammatic approach based on Wick's theorem [8], and 2) the functional derivative approach [9,10,23]. The former, for example, is used in the so-called electron propagator theory for molecular systems [26], where the screening among electrons is weak, and a many-body perturbation theory with respect to the bare Coulomb interaction up to the second or third order often gives very accurate descriptions as recently demonstrated by Ortiz and coworkers (e.g., [27] and references therein). For extended systems with stronger screening, the second approach, in the form of a set of coupled equations (Hedin's equations), is more appropriate, and is the subject of this work.

Hedin's equations, as first formulated in a complete form by Hedin [10] using the functional derivative technique, link the self-energy with the dynamically screened Coulomb interaction W , which, when calculated perturbatively, leads to a many-body perturbation expansion of the self-energy with respect to W instead of v . The derivation of Hedin's equations can be found in Refs. [9,23,24], and here we just give the final expressions with short explanations on the important quantities that are involved in the equations.

Using the short-hand notation $1 \equiv (\mathbf{x}_1, t_1)$, Hedin's equations read

$$\begin{aligned} \Sigma(1,2) &= i \int d(34) G(1,3) W(4,1) \Gamma(3,2,4) \\ \Gamma(1,2,3) &= \delta(1,2) \delta(2,3) + \int d(4567) \frac{\delta \Sigma(1,2)}{\delta G(4,5)} G(4,6) \\ &\quad \times G(7,5) \Gamma(6,7,3) \end{aligned}$$

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

$$P(1,2) = -i \int d(34) G(1,3) \Gamma(3,4,2) G(4,1^+) \quad (17)$$

which, together with the equation of motion for Green's function (Eq. (10)), form a close set of equations. In Eq. (17), P is the polarization function,

$$P(1,2) \equiv \frac{\delta \rho(1)}{\delta V(2)} \quad (18)$$

which describes the response of the electron density (ρ) with respect to the total effective potential $V = V_H + \phi$, with ϕ being the external perturbation. W is the screened Coulomb interaction defined as

$$W(1,2) = \int d(3) \varepsilon^{-1}(1,3) v(3,2) \quad (19)$$

with the inverse dielectric function ε^{-1} defined as

$$\varepsilon^{-1}(1,2) \equiv \frac{\delta V(1)}{\delta \phi(2)} \quad (20)$$

The dielectric function $\varepsilon(1,2)$ is related to the polarization function by

$$\varepsilon(1,2) = \delta(1,2) - \int d(3) v(1,3) P(3,2) \quad (21)$$

$\Gamma(1,2,3)$ is the vertex function defined as

$$\Gamma(1,2,3) \equiv -\frac{\delta G^{-1}(1,2)}{\delta V(3)} \quad (22)$$

Hedin's equations are a set of complicated integro-differential equations, which are not solvable even for the simplest systems like the homogeneous electron gas [9]. The main use of Hedin's equations is to serve as the starting point for a many-body perturbation theory in terms of W . Since the screening, the most important electron-electron interaction effect for many extended systems is a built-in feature of W , and even a low-order expansion of the self-energy with respect to W is hopeful to give reasonable description. In practice, the most widely used approach to solve Hedin's equations is the so-called GW approximation, which can be obtained by taking the zero-order approximation for the vertex function $\Gamma(1,2,3) \simeq \delta(1,2) \delta(2,3)$. In this case, the polarization function is simplified to be the product of two Green's functions,

$$P(1,2) = -i G(1,2) G(2,1^+) \quad (23)$$

which is often called random-phase approximation (RPA) for the historical reason. The self-energy now becomes a simple product of the Green's function G and screened Coulomb interaction W

$$\Sigma(1,2) = iG(1,2^+)W(2,1) \quad (24)$$

The positive infinitesimal in the equation above is needed for the bare exchange self-energy as a result of the Feynman's rules [8], $G(\mathbf{x}_1 t_1; \mathbf{x}_2 t_2 = t_1) \equiv G(\mathbf{x}_1 t_1; \mathbf{x}_2 t_1^+)$ [11,24]. When represented in the frequency domain, the GW self-energy reads

$$\Sigma(\mathbf{x}, \mathbf{x}'; \omega) = \frac{i}{2\pi} \int d\omega' e^{i\omega' \delta} G(\mathbf{x}, \mathbf{x}'; \omega + \omega') W(\mathbf{x}', \mathbf{x}; \omega') \quad (25)$$

3 GW in practice: Numerical techniques

3.1 The G_0W_0 approach

Even within the GW approximation Hedin's equations are still highly complicated, as illustrated in Fig. 1, requiring a self-consistent calculation of the self-energy from QP energies and wavefunctions as solutions of Dyson's equation. The latter is mathematically cumbersome to tackle due to the nonhermiticity of the self-energy. In practice, further approximations are often introduced. The most widely used implementation of the GW approximation is the so-called G_0W_0 or one-shot GW approach, in which both G and W are calculated based on eigen-energies ϵ_n and eigen-functions ψ_n of some non-interacting reference system,

$$\left[-\frac{1}{2} \nabla^2 + V_{\text{ext}}(\mathbf{x}) + V_{\text{H}}(\mathbf{x}) + V_{\text{xc}}(\mathbf{x}) \right] \psi_n(\mathbf{x}) = \epsilon_n \psi_n(\mathbf{x}) \quad (26)$$

where $V_{\text{xc}}(\mathbf{x})$ is the Kohn-Sham exchange-correlation potential, often in LDA or GGA. G_0 now reads

$$G_0(\mathbf{x}, \mathbf{x}'; \omega) = \sum_n \frac{\psi_n(\mathbf{x}) \psi_n^*(\mathbf{x}')}{\omega - \tilde{\epsilon}_n} \quad (27)$$

where $\tilde{\epsilon}_n \equiv \epsilon_n + i\eta \text{sgn}(\epsilon_F - \epsilon_n)$. The polarization function can then be calculated as

$$\begin{aligned} P_0(\mathbf{x}, \mathbf{x}'; \omega) &= -\frac{i}{2\pi} \int G_0(\mathbf{x}, \mathbf{x}'; \omega + \omega') G_0(\mathbf{x}', \mathbf{x}; \omega') d\omega' \\ &= \sum_{n,m} f_n (1 - f_m) \psi_n(\mathbf{x}) \psi_m^*(\mathbf{x}) \psi_n^*(\mathbf{x}') \psi_m(\mathbf{x}') \\ &\quad \times \left(\frac{1}{\omega - \epsilon_m + \epsilon_n + i\eta} - \frac{1}{\omega + \epsilon_m - \epsilon_n - i\eta} \right) \\ &\equiv \sum_{n,m} F_{nm}(\omega) \psi_n(\mathbf{x}) \psi_m^*(\mathbf{x}) \psi_n^*(\mathbf{x}') \psi_m(\mathbf{x}') \end{aligned} \quad (28)$$

where f_n is the occupation number of the n -th state, and $F_{nm}(\omega)$ denotes the occupation and frequency dependent factor. Using P_0 , the RPA screened Coulomb interaction can be obtained

$$\begin{aligned} W_0(\mathbf{x}, \mathbf{x}'; \omega) &= \int d\mathbf{x}'' \varepsilon^{-1}(\mathbf{x}, \mathbf{x}''; \omega) v(\mathbf{x}'', \mathbf{x}') \\ \varepsilon(\mathbf{x}, \mathbf{x}'; \omega) &= \delta(\mathbf{x}, \mathbf{x}') - \int d\mathbf{x}'' v(\mathbf{x}, \mathbf{x}'') P_0(\mathbf{x}'', \mathbf{x}'; \omega) \end{aligned} \quad (29)$$

It is practically common and convenient to decompose the self-energy into exchange and correlation terms by defining

$$W_0^c(\mathbf{x}, \mathbf{x}'; \omega) = W_0(\mathbf{x}, \mathbf{x}'; \omega) - v(\mathbf{x}, \mathbf{x}') \quad (30)$$

The exchange part, which is just given by the Hartree-Fock exchange potential, reads

$$\begin{aligned} \Sigma^x(\mathbf{x}, \mathbf{x}') &= \frac{i}{2\pi} \int G_0(\mathbf{x}, \mathbf{x}'; \omega') v(\mathbf{x}', \mathbf{x}) e^{i\omega' \eta} d\omega' \\ &= -\sum_n f_n \psi_n(\mathbf{x}) v(\mathbf{x}', \mathbf{x}) \psi_n^*(\mathbf{x}') \end{aligned} \quad (31)$$

The GW correlation self-energy is obtained from the frequency integral

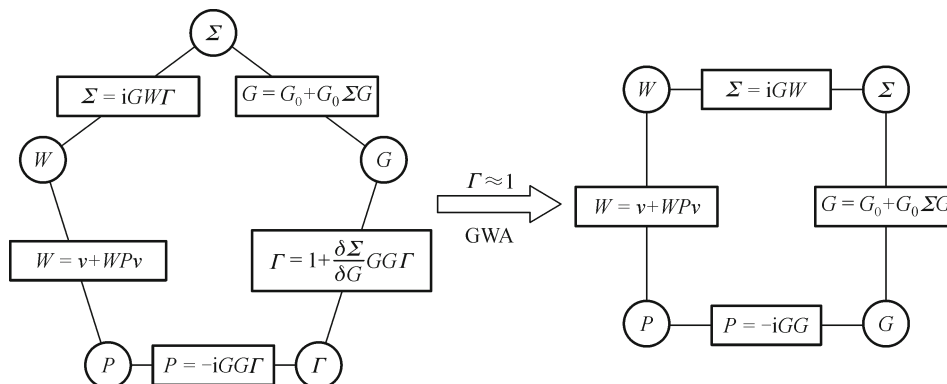


Figure 1 Schematic illustration of Hedin's equations in the full form (left) and within the GW approximation (right)

$$\Sigma^c(\mathbf{x}, \mathbf{x}'; \omega) = \frac{i}{2\pi} \int G_0(\mathbf{x}, \mathbf{x}'; \omega + \omega') W_0^c(\mathbf{x}', \mathbf{x}; \omega') d\omega' \quad (32)$$

The QP energies ε_n are calculated by the first order perturbation theory, treating $\delta\Sigma \equiv \Sigma - V_{xc}$ as the perturbation

$$\varepsilon_n = \varepsilon_n + Z_n(\varepsilon_n) \Re \langle \psi_n | \Sigma(\varepsilon_n) - V_{xc} | \psi_n \rangle \quad (33)$$

where Z_n is the QP renormalization factor,

$$Z_n(E) = \left[1 - \left(\frac{\partial}{\partial \omega} \Re \langle \psi_n | \Sigma(\omega) | \psi_n \rangle \right)_{\omega=E} \right]^{-1} \quad (34)$$

accounting for the frequency dependency of Σ . For *sp*-semiconductors it has been demonstrated that further improvement can be obtained by introducing partial self-consistency in the so-called GW_0 approach, in which the energies used for the calculation of the Green's function are updated by QP energies with fixed W_0 [22,28]. The GW_0 can be implemented with little computational overhead if the intermediate quantities used for Σ^c are stored during the calculation¹⁾.

3.2 The G_0W_0 equations in the matrix form

To put the GW approach in action, the first step is to convert the equations above into the matrix form. For that purpose we need a set of basis functions that are able to represent the products of any two Kohn-Sham wave functions accurately. Since we are mainly concerned with extended systems, we will include explicitly the dependence on wave-vectors (\mathbf{k}) that characterize the translational invariance of the system in the equations from now on unless stated otherwise. We use Ω to denote the volume of the unit cell. The difference between two \mathbf{k} vectors is denoted as \mathbf{q} . We first start with a general basis set, denoted as $\{\chi_i^{\mathbf{q}}(\mathbf{x})\}$. More technical details on using a particular basis will be discussed in the next subsection. To simplify the notation, we assume $\{\chi_i^{\mathbf{q}}(\mathbf{x})\}$ with the same \mathbf{q} are orthonormal, although non-orthonormal basis functions can also be used. In addition, we assume that the system under study is spin-unpolarized so that the spin index is dropped.

The most important requirement for the basis functions used in G_0W_0 is to represent the products of two Kohn-Sham wave functions accurately

$$\psi_{n\mathbf{k}}(\mathbf{r}) \psi_{m\mathbf{k}-\mathbf{q}}^*(\mathbf{r}) = \sum_i M_{nm}^i(\mathbf{k}, \mathbf{q}) \chi_i^{\mathbf{q}}(\mathbf{r}) \quad (35)$$

where $M_{nm}^i(\mathbf{k}, \mathbf{q})$ are expansion coefficients given by

$$M_{nm}^i(\mathbf{k}, \mathbf{q}) \equiv \int_V [\chi_i^{\mathbf{q}}(\mathbf{r}) \psi_{n\mathbf{k}-\mathbf{q}}(\mathbf{r})]^* \psi_{n\mathbf{k}}(\mathbf{r}) d^3\mathbf{r} \quad (36)$$

V is the total volume of the system, which is related to the volume of the unit cell by $V = N_k \Omega$ with N_k being the number of \mathbf{k} -points. We denote this basis the *product basis* to distinguish it from the basis that is used to expand single-particle wave functions. In quantum chemistry, it is also often called the ‘‘auxiliary’’ basis. Any two-point function, $\mathbf{g}(\mathbf{r}, \mathbf{r}')$, such as the polarizability $P_0(\mathbf{r}, \mathbf{r}'; \omega)$, the bare Coulomb interaction $v(\mathbf{r}, \mathbf{r}')$, the dielectric function $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$, and the screened Coulomb interaction $W(\mathbf{r}, \mathbf{r}'; \omega)$, will be expanded in terms of the product basis:

$$\mathbf{g}(\mathbf{r}, \mathbf{r}') = \sum_{\mathbf{q}} \sum_{ij}^{\text{BZ}} \mathbf{g}_{ij}(\mathbf{q}) \chi_i^{\mathbf{q}}(\mathbf{r}) [\chi_j^{\mathbf{q}}(\mathbf{r}')]^* \quad (37)$$

and the expansion coefficients are determined by

$$\mathbf{g}_{ij}(\mathbf{q}) \equiv \int_V \int_V [\chi_i^{\mathbf{q}}(\mathbf{r})]^* \mathbf{g}(\mathbf{r}, \mathbf{r}') \chi_j^{\mathbf{q}}(\mathbf{r}') \quad (38)$$

The diagonal elements of the bare (Fock) exchange self-energy can be written in the product basis as

$$\begin{aligned} \Sigma_{n\mathbf{k}}^{\mathbf{x}} &\equiv \int_V d\mathbf{r} \int_V d\mathbf{r}' [\psi_{n\mathbf{k}}(\mathbf{r})]^* \Sigma^{\mathbf{x}}(\mathbf{r}, \mathbf{r}') \psi_{n\mathbf{k}}(\mathbf{r}') \\ &= - \sum_{\mathbf{q}} \sum_{ij}^{\text{BZ}} v_{ij}(\mathbf{q}) \sum_m^{\text{occ}} [M_{nm}^i(\mathbf{k}, \mathbf{q})]^* M_{nm}^j(\mathbf{k}, \mathbf{q}) \end{aligned} \quad (39)$$

The matrix elements of the polarizability (Eq. (28)) can be written as

$$\begin{aligned} P_{ij}(\mathbf{q}, \omega) &\equiv \int_V \int_V [\chi_i^{\mathbf{q}}(\mathbf{r})]^* P(\mathbf{r}, \mathbf{r}'; \omega) \chi_j^{\mathbf{q}}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ &= 2 \sum_{\mathbf{k}} \sum_{n,m}^{\text{BZ}} F_{nm}(\mathbf{k}, \mathbf{q}; \omega) M_{nm}^i(\mathbf{k}, \mathbf{q}) [M_{nm}^j(\mathbf{k}, \mathbf{q})]^* \end{aligned} \quad (40)$$

where the factor of 2 is from the spin degeneracy, and

$$\begin{aligned} F_{nm}(\mathbf{k}, \mathbf{q}; \omega) &\equiv f_{n\mathbf{k}}(1 - f_{m\mathbf{k}-\mathbf{q}}) \left[\frac{1}{\omega - \omega_{n\mathbf{k}, m\mathbf{k}-\mathbf{q}} + i\eta} - \frac{1}{\omega + \omega_{n\mathbf{k}, m\mathbf{k}-\mathbf{q}} - i\eta} \right] \end{aligned} \quad (41)$$

with $\omega_{n\mathbf{k}, m\mathbf{k}-\mathbf{q}} \equiv \varepsilon_{m\mathbf{k}-\mathbf{q}} - \varepsilon_{n\mathbf{k}}$. Due to the singularity of the bare Coulomb interaction in reciprocal space as \mathbf{q} goes to zero, the dielectric function at $\mathbf{q} \rightarrow 0$ has to be treated carefully. Mathematically, it is more convenient to use the symmetrized form of the dielectric function [29], which, in the matrix form, can be obtained from $P_{ij}(\mathbf{q}, \omega)$ by

1) Jiang, H.; Gomez-Abal, R. I.; Li, X.; Meisenbichler, C.; Ambrosch-Draxl, C.; Scheffler, M. ‘‘Phi-gap: A green-function code based on augmented planewaves’’

$$\varepsilon_{ij}(\mathbf{q}, \omega) = \delta_{ij} - \sum_{lm} v_{il}^2(\mathbf{q}) P_{lm}(\mathbf{q}, \omega) v_{mj}^2(\mathbf{q}) \quad (42)$$

The correlation term of the screened Coulomb interaction can then be calculated through

$$W_{ij}^c(\mathbf{q}, \omega) = \sum_{lm} v_{il}^2(\mathbf{q}) [\varepsilon_{lm}^{-1}(\mathbf{q}, \omega) - \delta_{lm}] v_{mj}^2(\mathbf{q}). \quad (43)$$

The diagonal matrix elements of the correlation self-energy can be written as

$$\begin{aligned} \Sigma_{nk}^c(\omega) &\equiv \int_V d\mathbf{r} \int_V d\mathbf{r}' [\psi_{nk}(\mathbf{r})]^* \Sigma^c(\mathbf{r}, \mathbf{r}'; \omega) \psi_{nk}(\mathbf{r}') \\ &= \sum_{\mathbf{q}} \sum_m \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{X_{nm}(\mathbf{k}, \mathbf{q}; \omega')}{\omega + \omega' - \tilde{\varepsilon}_{m\mathbf{k}-\mathbf{q}}} \end{aligned} \quad (44)$$

where the auxiliary quantity $X_{nm}(\mathbf{k}, \mathbf{q}; \omega)$ is defined as

$$X_{nm}(\mathbf{k}, \mathbf{q}; \omega) \equiv \sum_{ij} [M_{nm}^i(\mathbf{k}, \mathbf{q})]^* W_{ij}^c(\mathbf{q}, \omega) M_{nm}^i(\mathbf{k}, \mathbf{q}) \quad (45)$$

We can see from the equations above that the main ingredients in the numerical implementation of G_0W_0 are 1) the bare Coulomb matrix, 2) the overlap integrals between the product basis functions and Kohn-Sham wave functions products, and 3) the summation of states in the calculation of P_0 and the self-energy. The treatment of 1) and 2) depends strongly on the nature of the product basis, and 3) is usually the most time-consuming part to calculate.

3.3 GW in different basis representations

The first "first-principles" implementation of the GW method by Hybertsen and Louie [12], and Godby et al. [30] used the planewave (PW) representation (also called the reciprocal space representation by many authors) in combination with the pseudopotential (PP) approximation. As we will see soon, the greatest advantage of the planewave representation is the simpleness in terms of implementation. In particular, since the planewave is the most popular basis for Kohn-Sham DFT for extended systems, the extension from DFT to GW is relatively straightforward. In addition, since the quality of the PW basis is controlled by a single parameter, the energy cutoff, the accuracy of the representation can be systematically monitored. Nevertheless, the limitation of the PW representation is also obvious. For systems with d - or f -states not very far way from the Fermi level, it is often necessary to treat these states as valence, but since these states are much more localized than sp states, the number of the planewaves required to represent them is usually quite large, leading to very heavy calculations. For systems with loose structure or a lot of open space such as weakly bonded molecular crystals and surfaces, the PW

representation is also not very efficient. There is another more technical drawback of PW: Since PWs are global basis functions, the matrices represented by PWs are usually far from sparse, which makes parallelization of the corresponding GW code much more difficult than those based on local basis functions.

To overcome the limitation of the PW basis functions, a lot of efforts have been invested in the past two decades to develop new techniques that are either more accurate or can treat larger systems than that the PW representation can afford. In the following, we present an overview on the implementations of GW with different basis functions.

3.3.1 The planewaves representation

When using the planewaves

$$\chi_i^{\mathbf{q}}(\mathbf{r}) \rightarrow \chi_{\mathbf{G}}^{\mathbf{q}}(\mathbf{r}) \equiv \frac{1}{\sqrt{V}} \exp[i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{r}] \quad (46)$$

as the product basis, then the integrals $M_{nm}^i(\mathbf{k}, \mathbf{q})$ become simply the Fourier transform of $\psi_{nk}(\mathbf{r}) \psi_{m\mathbf{k}-\mathbf{q}}^*(\mathbf{r})$. In addition, since $\psi_{nk}(\mathbf{r})$ is also represented by plane waves,

$$\psi_{nk} = \sum_{\mathbf{G}} c_{nk; \mathbf{G}} \chi_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r}) \quad (47)$$

one obtains

$$M_{nm}^{\mathbf{G}}(\mathbf{k}, \mathbf{q}) = V^{-1/2} \sum_{\mathbf{G}'} c_{nk; \mathbf{G}'} C_{m\mathbf{k}-\mathbf{q}; \mathbf{G}' - \mathbf{G}}^* \quad (48)$$

An even more favorable feature of the PW basis is that the bare Coulomb matrix is diagonal in the PW representation

$$v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \frac{4\pi}{|\mathbf{q} + \mathbf{G}|^2} \delta_{\mathbf{G}, \mathbf{G}'} \quad (49)$$

The symmetrized dielectric function now reads

$$\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \delta_{\mathbf{G}\mathbf{G}'} - \frac{4\pi}{|\mathbf{q} + \mathbf{G}| |\mathbf{q} + \mathbf{G}'|} P_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) \quad (50)$$

When \mathbf{q} is at the Γ point, i.e., $\mathbf{q} = 0$, the second term in Eq. (50) is divergent when $\mathbf{G} = 0$ and/or $\mathbf{G}' = 0$. However, this singularity can be easily removed by expanding $P_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega)$ around $\mathbf{q} = 0$ when $\mathbf{G} = 0$ and/or $\mathbf{G}' = 0$, which gives the lowest order term $\propto |\mathbf{q}|^2$ (when both $\mathbf{G} = 0$ and $\mathbf{G}' = 0$) or $\propto |\mathbf{q}|$ ($\mathbf{G} = 0$ or $\mathbf{G}' = 0$).

3.3.2 The space-time approach

A variant of the PW representation is the space-time approach developed by R. Godby and coworkers [31,32] The basic ideas are as follow.

(i) Since both the RPA polarization function and the GW

self-energy are simple products in the space and time domain,

$$P_0(\mathbf{r}, \mathbf{r}'; \tau) = -iG_0(\mathbf{r}, \mathbf{r}'; \tau)G_0(\mathbf{r}', \mathbf{r}; -\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}'; \tau) = iG(\mathbf{r}, \mathbf{r}'; \tau)W(\mathbf{r}', \mathbf{r}; \tau) \quad (51)$$

it is computationally more efficient to calculate them directly in the space-time representation instead of in the plane-wave representation.

(ii) On the other hand, the evaluation of the dielectric function ε and the screened Coulomb interaction (W) involves the expensive integration in the real space. It is more convenient to calculate them in the reciprocal space to take advantage of the diagonality of the bare Coulomb interaction. The space-time representation and the plane-waves-frequency representation can be converted to each other by the fast-Fourier transform technique.

(iii) To avoid tackling the complicated structures that Σ , W and G have along the real frequency/time, it is more efficient to construct these functions first on the imaginary time/frequency. Self-energies along the real frequency axis can be obtained by the analytic continuation technique, thanks to the analyticity of these functions in the complex frequency domain.

It is important to bear in mind that since in the space-time approach, quantities like G , P , W and Σ need to be represented in both real-space (usually on a uniform grid) and the reciprocal space, the effective potential of the system under study has to be quite smooth, so that it is necessary to use the pseudo-potential approximation, as in the standard plane-waves approach.

3.3.3 The local basis representation

A completely different framework for the implementation of the GW method is to use local atomic-like basis. The latter can be analytic Gaussian-type orbitals (GTO) [33,34] that is the de facto standard in molecular quantum chemistry community [35], or numerical atomic orbitals [36]. The general matrix formulation in the preceding subsection is based on an orthogonal basis representation. When using the local atomic-like (numerical or analytic) basis functions, basis functions are in general not orthogonal, for which particular attention is needed to obtain a robust and efficient implementation.

Local basis functions written in the Bloch form reads

$$\chi_\alpha^{\mathbf{q}}(\mathbf{r}) = \frac{1}{N_c^{1/2}} \sum_{\mathbf{R}} e^{i\mathbf{q} \cdot (\mathbf{R} + \tau_\alpha)} \phi_\alpha(\mathbf{r} - \mathbf{R} - \tau_\alpha) \quad (52)$$

where $\phi_\alpha(\mathbf{r})$ are atomic-like functions centered at τ_α . We use the index α to denote both the positions of the atoms and the atomic quantum numbers that characterize the atomic basis

functions. Now we consider a general two-point function $g(\mathbf{r}, \mathbf{r}')$, and define the matrix elements

$$[g]_{\alpha\beta}(\mathbf{q}) \equiv \int_V d\mathbf{r} \int_V d\mathbf{r}' \chi_\alpha^{\mathbf{q}*}(\mathbf{r}) g(\mathbf{r}, \mathbf{r}') \chi_\beta^{\mathbf{q}}(\mathbf{r}') \quad (53)$$

Then $g(\mathbf{r}, \mathbf{r}')$ can be expanded by the local basis functions as

$$g(\mathbf{r}, \mathbf{r}') = \sum_{\mathbf{q}} \sum_{\alpha, \beta} \chi_\alpha^{\mathbf{q}}(\mathbf{r}) \langle g \rangle_{\alpha\beta}(\mathbf{q}) \chi_\beta^{\mathbf{q}*}(\mathbf{r}') \quad (54)$$

The matrix $\langle g \rangle$ is related to $[g]$ by (in the matrix form)

$$\langle g \rangle(\mathbf{q}) = \mathbf{S}^{-1}(\mathbf{q}) [g](\mathbf{q}) \mathbf{S}^{-1}(\mathbf{q}) \quad (55)$$

where $\mathbf{S}(\mathbf{q})$ is the overlap matrix

$$S_{\alpha\beta}(\mathbf{q}) \equiv \int_V d\mathbf{r} \chi_\alpha^{\mathbf{q}}(\mathbf{r}) \chi_\beta^{\mathbf{q}*}(\mathbf{r}) \quad (56)$$

To obtain the GW equations in the non-orthogonal basis representation, we can introduce the orthogonalized basis corresponding to $\chi_\alpha^{\mathbf{q}}(\mathbf{r})$

$$|\tilde{\chi}_i^{\mathbf{q}}\rangle = \sum_{\alpha} |\chi_\alpha^{\mathbf{q}}\rangle S_{\alpha i}^{-1/2}(\mathbf{q}) \quad (57)$$

The orthogonal and non-orthogonal representations are related by

$$[\tilde{g}](\mathbf{q}) = \mathbf{S}^{-1/2}(\mathbf{q}) [g](\mathbf{q}) \mathbf{S}^{-1/2}(\mathbf{q}) \quad (58)$$

Using the relation above, one can easily convert the equations in Section 3.2 into their counter-parts in the non-orthogonal basis representation.

A relatively simpler scheme to handle the nonorthogonality of atomic basis functions is to use the orthogonalized basis from the very beginning [34,36]. The basic ingredients such as the bare Coulomb matrix v_{ij} and the product expansion matrix M_{nm}^i are first calculated in the original non-orthogonal basis, and then are transformed into the orthogonalized form. Then all subsequent treatments are essentially same as that in the orthogonal basis representation. In addition, by using this pre-orthogonalized basis, the linear-dependency of the non-orthogonal basis functions, usually occurring between basis functions centered on different atoms, can be removed from the very beginning by dropping those eigenvectors of the overlap matrix that have vanishing eigen-values [34].

3.3.4 The mixed basis representation

One of the main concerns over using the plane-wave representation or the space-time approach is that the pseudopotential approximation is necessary. In principles, the error introduced by the use of the PP approximation can be controlled if all relevant atomic states are treated as “valence,” but in practice the situations are more complicated. In

KS-DFT, the accuracy of using the PP relies on the validity of two approximations: 1) the frozen-core approximation, i.e., that the states treated as "core" are chemically inert, and 2) the linear approximation for the exchange-correlation potential as the functional of the electron density. In the *GW* method, there is one more aspect of using the PP approximation: since the *GW* self-energy depends on wave functions instead of electron density, the use of the pseudo-wave functions can also have significant influences on the accuracy of the *GW* [37–39]: 1) the use of the PP approximation has much stronger effects on the accuracy of the *GW* results than it does on KS-DFT calculations; and 2) for many elements, it is often necessary to use the pseudopotentials that treat shallow semi-core states as "valence," and therefore they are much harder than those used in KS-DFT.

To avoid the difficulty of using the PP approach, and in the meanwhile maintain the advantage of the planewave basis that the accuracy can be systematically improved, the full-potential augmented basis approaches have been developed for the implementation of *GW* [40–43] (see footnote on Page 6). We will use the full-potential linearized-planewaves (FP-LAPW) as an example to describe the main features of this type of approaches (see footnote on Page 6). In the FP-LAPW approach, the space in the unit cell is partitioned into non-overlapping muffin-tin (MT) spheres, centered around each atom (indexed by α , positioned at τ_α in the unit cell), and the interstitial (IS) region. The LAPW basis is given in the IS region as plane-waves

$$\phi_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \quad \mathbf{r} \in \text{IS} \quad (59)$$

In the MT spheres, it is represented by atomic-like wave functions

$$\begin{aligned} \phi_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r}) = & \sum_{lm} [A_{alm}(\mathbf{k} + \mathbf{G})u_{al}(r^\alpha) \\ & + B_{alm}(\mathbf{k} + \mathbf{G})\dot{u}_{al}(r^\alpha)] Y_{lm}(\hat{\mathbf{r}}^\alpha), \quad r^\alpha < R_{\text{MT}}^\alpha \quad (60) \end{aligned}$$

where $\mathbf{r}^\alpha \equiv \mathbf{r} - \tau_\alpha$ and R_{MT}^α is the radius of α -th MT sphere. $u_{al}(r^\alpha)$ are the solutions of the radial Schrödinger equation in the spherical potential of the respective MT sphere, taken at a l -dependent reference energy E_l , and $\dot{u}_{al}(r^\alpha)$ is the first order derivative of $u_{al}(r^\alpha)$ with respect to the energy at E_l . The augmentation coefficients, $A_{alm}(\mathbf{k} + \mathbf{G})$ and $B_{alm}(\mathbf{k} + \mathbf{G})$, are determined from the continuity of the basis functions and their first derivatives at the MT sphere boundary.

To implement *GW* in the FP-LAPW framework, the product basis with similar mixed features should be used to represent the products of KS wave functions accurately, hence termed as the mixed basis (MB). The MB functions can be constructed in the following way: Inside the MT sphere of

atom α , the basis functions are atomic like, similar to the product basis originally proposed by Aryasetiawan and Gunnarsson [40].

$$\chi_i^{\mathbf{k}}(\mathbf{r}) = \frac{1}{N_c^{1/2}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\tau_\alpha)} v_{\alpha NL}(|\mathbf{r}-\tau_\alpha-\mathbf{R}|) Y_{LM}(\hat{\mathbf{r}}) \quad (61)$$

The radial functions $v_{\alpha NL}(r^\alpha)$ are often constructed from the products of the radial wave functions used in the LAPW basis (see Eq. (60)). Since the number of the latter is quite big, and in addition, they are not fully linearly independent, the following procedure is often used to construct an optimal set of orthonormal radial functions:

- To reduce the number of product functions, only $u_{al}(r^\alpha)$'s with $l \leq l_{\text{max}}^{\text{MB}}$ are considered; in addition, $\dot{u}_{al}(\mathbf{r})$'s are not taken into account as they are typically one order smaller than $u_{al}(\mathbf{r})$ [44,45].

- For each L , all the products of two radial functions $u_{al}(r^\alpha)u_{al'}(r^\alpha)$ that fulfill the triangular condition $|l-l'| \leq L \leq l+l'$ are considered.

- The overlap matrix between this set of product radial functions

$$\mathbb{O}_{l'l'} = \int_0^{R_{\text{MT}}^\alpha} u_{al}(\mathbf{r})u_{al'}(\mathbf{r})u_{al_1}(\mathbf{r})u_{al'_1}(\mathbf{r})\mathbf{r}^2 d\mathbf{r} \quad (62)$$

is diagonalized, yielding the corresponding set of eigenvalues, λ_N^{MB} , and eigenvectors, $\{c_{ll',N}\}$.

- Eigenvectors corresponding to eigenvalues smaller than a certain tolerance, $\lambda_{\text{min}}^{\text{MB}}$ (typically about 10^{-4}), are assumed to be linearly dependent and discarded [38]; The remaining eigenvectors, after normalization, form the radial functions

$$v_{\alpha NL}(r^\alpha) = \sum_{l'l'} c_{ll',N} u_{al}(r^\alpha) u_{al'}(r^\alpha) \quad (63)$$

In the interstitial region, the product basis functions are constructed as ortho-normalized interstitial planewaves (IPW's),

$$P_i^{\mathbf{q}}(\mathbf{r}) \equiv \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} \frac{S_{G_i}}{\lambda_i^{\text{IS}}} e^{i(\mathbf{G}+\mathbf{q})\cdot\mathbf{r}} \theta_{\text{IS}}(\mathbf{r}) \quad (64)$$

where

$$\theta_{\text{IS}}(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in \text{IS}, \\ 0 & \text{otherwise.} \end{cases} \quad (65)$$

S_{G_i} and λ_i^{IS} are eigenvectors and eigenvalues of the overlap matrix between IPW's

$$\mathbb{O}_{\mathbf{G}\mathbf{G}'} = \frac{1}{\Omega} \int_{\Omega} \theta_{\text{IS}}(\mathbf{r}) e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}} d^3\mathbf{r} \equiv \frac{1}{\Omega} \mathbb{I}_{\mathbf{G}-\mathbf{G}'} \quad (66)$$

where $\mathbb{I}_{\mathbf{G}}$ is given by

$$\mathbb{I}_{\mathbf{G}} = \Omega \delta_{\mathbf{G},0} - \sum_{\alpha} V_{\text{MT}}^{\alpha} e^{i\mathbf{G} \cdot \mathbf{r}_{\alpha}} \mathcal{I}(\mathbf{G} R_{\text{MT}}^{\alpha}) \quad (67)$$

with $V_{\text{MT}}^{\alpha} = \frac{4\pi}{3}(R_{\text{MT}}^{\alpha})^3$ being the volume of MT sphere α , and

$$\mathcal{I}(x) \equiv \frac{3(\sin x - x \cos x)}{x^3} \quad (68)$$

To summarize, the mixed basis set is given by

$$\{\chi_j^{\mathbf{q}}(\mathbf{r})\} \equiv \{\gamma_{\alpha NLM}^{\mathbf{q}}(\mathbf{r}), P_i^{\mathbf{q}}(\mathbf{r})\} \quad (69)$$

The main ingredients of the implementation using the mixed basis is the construction of the bare Coulomb matrix $v_{ij}(\mathbf{q})$ and the evaluation of the production expansion coefficients $M_{nm}^i(\mathbf{k}, \mathbf{q})$. More details can be found in the footnoted reference (see the footnote on Page 6).

3.4 Frequency dependence

How to treat the frequency dependence of the dielectric function has significant influences on the accuracy of the results as well as the efficiency of the implementation. Various techniques or approximations have been developed and roughly they fall into the following categories:

- static approximations
- generalized plasmon models
- imaginary frequency plus analytic continuation approach
- the Hilbert transform approach
- the contour deformation approach

In the following we will discuss the essence of each scheme in turn.

3.4.1 Static approximations

The simplest approximation concerning the treatment of frequency is to neglect the frequency dependence at all. The first static approximation to GW is the so-called static Coulomb-hole and screened-exchange (COHSEX) approximation proposed by Hedin [10]. The real part of the self-energy can be written as

$$\begin{aligned} \text{Re}\Sigma(\mathbf{r}, \mathbf{r}'; \omega) &= - \sum_{nk}^{\text{occ}} \psi_{nk}(\mathbf{r}) \psi_{nk}^*(\mathbf{r}') \Re \mathcal{W}(\mathbf{r}', \mathbf{r}; \omega - \epsilon_{nk}) \\ &\quad - \sum_{nk} \psi_{nk}(\mathbf{r}) \psi_{nk}^*(\mathbf{r}') \frac{1}{\pi} \mathcal{P} \int_0^{\infty} d\omega' \frac{\Im W_c(\mathbf{r}', \mathbf{r}; \omega)}{\omega - \epsilon_{nk} - \omega'} \end{aligned} \quad (70)$$

where \mathcal{P} represents taking the Cauchy principle value. The static COHSEX approximation can be obtained by setting $\omega - \epsilon_{nk} = 0$ in Eq. (70), which gives

$$\Sigma^{\text{COHSEX}}(\mathbf{r}, \mathbf{r}') = \Sigma^{\text{COH}}(\mathbf{r}, \mathbf{r}') + \Sigma^{\text{SEX}}(\mathbf{r}, \mathbf{r}')$$

$$\Sigma^{\text{COH}}(\mathbf{r}, \mathbf{r}') = \frac{1}{2} \delta(\mathbf{r} - \mathbf{r}') [W(\mathbf{r}, \mathbf{r}'; 0) - v(\mathbf{r} - \mathbf{r}')] \quad (71)$$

$$\Sigma^{\text{SEX}}(\mathbf{r}, \mathbf{r}') = - \sum_{nk}^{\text{occ}} \psi_{nk}(\mathbf{r}) \psi_{nk}^*(\mathbf{r}') W(\mathbf{r}, \mathbf{r}'; 0)$$

The static COHSEX tends to overestimate the band gaps of many semiconductors [46]. On the other hand, since the COHSEX self-energy is hermitian, it can be calculated self-consistently, and the resultant single-particle energies and wave functions can be used as the starting point for a full G_0W_0 calculation. This self-consistent COHSEX based G_0W_0 approach has been used recently for a series of systems with remarkable success [15,47–50].

An approach closely related to the static COHSEX approximation is the model GW approach first proposed by Gygi and Baldereschi [51], in which the self-energy correction with respect to the LDA exchange-correlation potential is approximated by

$$\delta\Sigma(\mathbf{r}, \mathbf{r}') = -\rho(\mathbf{r}, \mathbf{r}') \delta W(\mathbf{r} - \mathbf{r}') \quad (72)$$

where $\rho(\mathbf{r}, \mathbf{r}')$ is the density matrix function, and $\delta W(\mathbf{r} - \mathbf{r}')$ is a model function for the screened Coulomb interaction correction, whose Fourier transform takes the form

$$\delta W(\mathbf{q}) = \frac{4\pi}{\Omega \mathbf{q}^2} [\epsilon_{\text{SC}}^{-1}(\mathbf{q}, \omega = 0) - \epsilon_{\text{M}}^{-1}(\mathbf{q}, \omega = 0)] \quad (73)$$

$\epsilon_{\text{SC}}^{-1}$ and ϵ_{M}^{-1} are the inverse dielectric functions for a semiconductor and a metal, respectively, both approximated by some model functions. This model GW approach was used by Massida and coworkers to investigate electronic band structure properties of transition metal oxides [52–54].

3.4.2 Generalized plasmon models

In the random-phase approximation to the polarization, the screening mainly comes from two types of excitations [11], the plasmon excitation and electron-hole excitations. The former corresponds to the collective excitation of all electrons with respect to fixed positive ions. It turns out that for self-energy, the plasmon excitation is dominant, and it is therefore physically plausible to approximate the polarization by retaining only the plasmon excitation, leading to the plasmon pole model. In the following, we will mainly use the planewaves representation for the discussion since various plasmon models originate from the homogeneous electron gas (HEG), for which the reciprocal space is the representation of choice.

For the homogeneous electron system, the plasmon pole model gives the polarization with a single pole [11,46]

$$\Im \varepsilon^{-1}(q, \omega) = A(q) \delta(\omega - \omega_p(q)) \quad (74)$$

where $\omega(q)$ is q -dependent plasmon frequency. For inhomogeneous systems, this plasmon pole model is usually generalized into the following form

$$\Im \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega) = A_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}) \delta(\omega - \tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})) \quad (75)$$

hence called generalized plasmon pole (GPP) model. Using the Kramers-Kronig's relation between the real and imaginary part of the inverse dielectric function [3]

$$\Re \varepsilon^{-1}(\mathbf{q}, \omega) = 1 + \frac{2}{\pi} \mathcal{P} \int_0^\infty d\omega' \frac{\omega' \Im \varepsilon^{-1}(\mathbf{q}, \omega')}{\omega'^2 - \omega^2} \quad (76)$$

one has

$$\Re \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega) = \delta_{\mathbf{G}\mathbf{G}'} + \frac{2}{\pi} \frac{\tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) A_{\mathbf{G}\mathbf{G}'}(\mathbf{q})}{\tilde{\omega}_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q}) - \omega^2} \quad (77)$$

For given \mathbf{G} , \mathbf{G}' and \mathbf{q} , there are two unknown parameters, $A_{\mathbf{G}, \mathbf{G}'}(\mathbf{q})$ and $\tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$ that need to be fixed.

Hybertsen-Louie (HL) model: In the Hybertsen-Louie's GPP model [12], the two parameters are determined by 1) the static dielectric function ($\omega = 0$); and 2) the f-sum rules [55]

$$\int_0^\infty \omega \Im \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega) d\omega = -\frac{\pi}{2} \Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q}) \quad (78)$$

where

$$\Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q}) \equiv \frac{4\pi(\mathbf{q} + \mathbf{G}) \cdot (\mathbf{q} + \mathbf{G}')}{|\mathbf{q} + \mathbf{G}| |\mathbf{q} + \mathbf{G}'|} \rho(\mathbf{G} - \mathbf{G}'). \quad (79)$$

The final expressions are

$$\begin{aligned} A_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) &= \frac{\pi}{2} [\delta_{\mathbf{G}\mathbf{G}'} - \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, 0)]^{1/2} |\Omega_{\mathbf{G}\mathbf{G}'}| \\ \tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) &= \frac{|\Omega_{\mathbf{G}\mathbf{G}'}|}{[\delta_{\mathbf{G}\mathbf{G}'} - \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, 0)]^{1/2}} \end{aligned} \quad (80)$$

Godby-Needs (GN) model: Godby and Needs [56] proposed a different way to determine the parameters in the GPP model: besides the requirement to give the static inverse dielectric function, the second condition they introduced is that the GPP model reproduce the inverse dielectric function at a chosen imaginary frequency, usually $\omega = i\omega_p$ with $\omega_p \equiv \sqrt{4\pi\rho(0)}$ being the classical Drude plasmon frequency. In this case, one obtains

$$\begin{aligned} A_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) &= \frac{\pi}{2} \omega_p^{1/2} [(\delta_{\mathbf{G}\mathbf{G}'} - \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, 0)) \\ &\times (\varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, 0) - \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, i\omega_p))]^{1/2} \end{aligned} \quad (81)$$

$$\tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \omega_p^{1/2} \left[\frac{\varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, 0) - \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, i\omega_p)}{\delta_{\mathbf{G}\mathbf{G}'} - \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, 0)} \right]^{1/2}$$

von der Linden-Horsch (vdLH) model: In the HL or GN GPP models, for each \mathbf{q} there are totally $2N_{\text{PW}}^2$ (N_{PW} : the number of planewaves) parameters to determine, which is not always numerically stable. In some cases, unphysical complex plasmon frequencies can arise for very large \mathbf{G} . To overcome this difficulty, von der Linden and Horsch (vdLH) [57] developed a different GPP model that depends on only 2 N_{PW} parameters, and is numerically robust. The first step in the construction of the vdLH model is to diagonalize the the symmetrized static dielectric function

$$\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, 0) = \sum_i U_{\mathbf{G}i}(\mathbf{q}) d_i(\mathbf{q}) U_{\mathbf{G}'i}^*(\mathbf{q}) \quad (82)$$

where $\{U_{\mathbf{G}i}\}$ and $\{d_i\}$ are eigenvectors and eigenvalues of $\varepsilon(\mathbf{q}, 0)$, respectively. The inverse dielectric function at arbitrary frequency is obtained by assuming that eigenvectors $\{U_{\mathbf{G}i}\}$ do not depend on the frequency, and the frequency dependence occurs only for eigenvalues.

$$\varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega) = \sum_i U_{\mathbf{G}i}(\mathbf{q}) d_i^{-1}(\mathbf{q}, \omega) U_{\mathbf{G}'i}^*(\mathbf{q}) \quad (83)$$

and it is assumed that $d_i^{-1}(\mathbf{q}, \omega)$ takes the plasmon-pole model like the expression

$$d_i^{-1}(\mathbf{q}, \omega) = 1 + \frac{z_i(\mathbf{q})}{\omega^2 - [\omega_i(\mathbf{q}) - i\eta]^2} \quad (84)$$

As in the HL GPP model, the parameters $z_i(\mathbf{q})$ and $\omega_i(\mathbf{q})$ are determined again from the static limit

$$d_i^{-1}(\mathbf{q}) = 1 - \frac{z_i(\mathbf{q})}{\omega_i^2(\mathbf{q})} \quad (85)$$

and the f-sum rules. After some algebraic formulation, one obtains

$$\begin{aligned} z_i(\mathbf{q}) &= \sum_{\mathbf{G}\mathbf{G}'} U_{\mathbf{G}'i}^*(\mathbf{q}) \Omega_{\mathbf{G}\mathbf{G}'}^2 U_{\mathbf{G}i}(\mathbf{q}) \\ \omega_i(\mathbf{q}) &= \left[\frac{z_i(\mathbf{q})}{1 - d_i^{-1}(\mathbf{q})} \right]^{1/2} \end{aligned} \quad (86)$$

Engel-Farid model: A further extension was made by G. E. Engel and B. Farid (EF) [58]: besides taking the f-sum rule into account, the exact behavior at large frequency limit is also taken into account. Based on the formal analysis, the inverse dielectric function at large frequency limit satisfies the following relation

$$\lim_{\omega \rightarrow \infty} \omega^2 \varepsilon^{-1}(\mathbf{r}, \mathbf{r}'; \omega) = -\frac{2}{\pi} \int_0^\infty \omega \Im \varepsilon^{-1}(\mathbf{r}, \mathbf{r}'; \omega) d\omega \quad (87)$$

In the reciprocal space representation and using the f-sum rules (Eq. (78)), one obtains

$$\lim_{\omega \rightarrow \infty} \omega^2 \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega) = \Omega_{\mathbf{G}, \mathbf{G}'}^2(\mathbf{q}) \equiv L_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}) \quad (88)$$

The EF GPP model for the inverse dielectric function is then defined in the planewave representation as (in the matrix form)

$$\varepsilon^{-1}(\mathbf{q}, \omega) = [\omega^2 L^{-1} + \varepsilon(\mathbf{q}, 0)]^{-1} \quad (89)$$

3.4.3 The imaginary frequency plus analytic continuation approach

A relatively simple approach to consider the full frequency dependence of the dielectric function and the correlation selfenergy is to compute them first on the imaginary frequency axis, $\omega = iu$, and then obtain the self-energy on the real axis by analytic continuation, hence termed as the IF + AC scheme. The polarization function is a smooth function of the imaginary frequency, so that only a small number of discrete frequency points are needed

$$P_{ij}(\mathbf{q}, iu) = 2 \sum_{\mathbf{k}}^{BZ} \sum_n^{\text{occ}} \sum_m^{\text{unocc}} \frac{-2(\varepsilon_{n\mathbf{k}} - \varepsilon_{n\mathbf{k}-\mathbf{q}})}{u'^2 + (\varepsilon_{n\mathbf{k}} - \varepsilon_{n\mathbf{k}-\mathbf{q}})^2} \times M_{nm}^i(\mathbf{k}, \mathbf{q}) [M_{nm}^j(\mathbf{k}, \mathbf{q})]^* \quad (90)$$

Making use of the inversion symmetry of W^c on the imaginary frequency axis, $W_{ij}^c(\mathbf{q}, iu) = W_{ij}^c(\mathbf{q}, -iu)$, the correlation self-energy for the imaginary frequency reads

$$\Sigma_{nk}^c(iu) = \sum_{\mathbf{q}}^{BZ} \sum_m \int_0^\infty \frac{du'}{2\pi} \frac{2(\varepsilon_{m\mathbf{k}-\mathbf{q}} - iu) X_{nm}(\mathbf{k}, \mathbf{q}; iu')}{u'^2 + (\varepsilon_{m\mathbf{k}-\mathbf{q}} - iu)^2} \quad (91)$$

The integrand in Eq. (91) is peaked around $u' = u$ for small $\varepsilon_{m\mathbf{k}-\mathbf{q}}$ such that a direct numerical integration is unstable. This can be avoided by adding and subtracting the following analytically integrable term [40]

$$\int_0^\infty \frac{du'}{2\pi} X_{nm}(\mathbf{k}, \mathbf{q}; iu) \frac{(\varepsilon_{m\mathbf{k}-\mathbf{q}} - iu)}{u'^2 + (\varepsilon_{m\mathbf{k}-\mathbf{q}} - iu)^2} = \frac{1}{2} \text{sgn}(\varepsilon_{m\mathbf{k}-\mathbf{q}}) X_{nm}(\mathbf{k}, \mathbf{q}; iu) \quad (92)$$

which gives

$$\Sigma_{nk}^c(iu) = \sum_{\mathbf{q}}^{BZ} \sum_m \left\{ \int_0^\infty \frac{du'}{2\pi} [X_{nm}(\mathbf{k}, \mathbf{q}; iu') - X_{nm}(\mathbf{k}, \mathbf{q}; iu)] \right.$$

$$\left. \times \frac{2(\varepsilon_{m\mathbf{k}-\mathbf{q}} - iu)}{u'^2 + (\varepsilon_{m\mathbf{k}-\mathbf{q}} - iu)^2} + \frac{1}{2} \text{sgn}(\varepsilon_{m\mathbf{k}-\mathbf{q}}) X_{nm}(\mathbf{k}, \mathbf{q}; iu) \right\} \quad (93)$$

The integrand is now a smooth function of u for any $\varepsilon_{m\mathbf{k}-\mathbf{q}}$, and therefore a standard Gaussian quadrature can be used. In practice, a double Gauss-Legendre quadrature [13,59] is often used, in which the semi-infinite integral is divided into two intervals, $[0, \omega_0]$ and $[\omega_0, \infty)$, and the integration in each interval is carried out by standard Gauss-Legendre quadrature. In most cases, the integration converges quite quickly with respect to the number of discrete frequency points when an appropriate ω_0 is chosen [59] (see footnote on Page 6).

Once the correlation self-energy matrix elements along the imaginary axis are calculated according to Eq. (93), they can be fitted by a function with N_p poles [31]

$$\Sigma_{nk}^c(iu) = \sum_p^{N_p} \frac{a_{p;nk}}{iu - b_{p;nk}} \quad (94)$$

where $a_{p;nk}$ and $b_{p;nk}$ are fitting parameters. Eq. (94) is then analytically continued onto the real frequency axis.

The IF + AC approach to treat the frequency dependence is relatively simple to implement and quite efficient in terms of the computational demand, but it has also such a disadvantage that the error of the analytic continuation cannot be controlled: Increasing the number of imaginary frequencies does not guarantee a convergence to more accurate real-frequency self-energies.

3.4.4 The Hilbert transform approach

To avoid the uncertainty related to the use of the IF + AC approach, one can treat everything along the real frequency directly. In this case, it is often advantageous to use the spectral representation of G , P , W and Σ^c [11,60].

Polarization matrix: In the general matrix form, the spectral representation of the polarization function reads

$$S_{ij}(\mathbf{q}, \omega) \equiv -\frac{1}{\pi} \Im P_{ij}(\mathbf{q}, \omega) \text{sgn}(\omega) = 2 \sum_{\mathbf{k}}^{BZ} \sum_n^{\text{occ}} \sum_m^{\text{unocc}} M_{nm}^i(\mathbf{k}, \mathbf{q}) [M_{nm}^j(\mathbf{k}, \mathbf{q})]^* \times \delta(|\omega| - \omega_{n\mathbf{k}, m\mathbf{k}-\mathbf{q}}) \text{sgn}(\omega) \quad (95)$$

Since only a small number of occupied-unoccupied pairs contribute for each frequency ω , $S_{ij}(\mathbf{q}, \omega)$ can be calculated quite efficiently. The full polarization function can then be calculated by the following Hilbert transform

$$P_{ij}(\mathbf{q}, \omega) = \int_{-\infty}^{\infty} d\omega' \frac{S_{ij}(\mathbf{q}, \omega')}{\omega - \omega' - i\eta \text{sgn}(\omega')} \quad (96)$$

Correlation self-energies Correlation self-energies on real-frequency can be calculated directly by using Eq. (44) [60]: one first calculates the auxiliary quantities $X_{nm}(\mathbf{k}, \omega)$ on a series of real frequencies, and the correlation self-energies are then calculated by a numerical integration over real frequency in terms of Eq. (44). This scheme is useful when the full frequency dependent self-energy is required. On the other hand, in G_0W_0 calculation, only the self-energy and its first-order derivative at the corresponding Kohn-Sham orbital energy are needed (See Eq. (33)). Following Ref. [60], first define the Hilbert-transformed screened Coulomb interaction

$$\tilde{W}_{ij}^{\pm}(\mathbf{q}, \omega) \equiv \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{W_{ij}^c(\mathbf{q}, \omega')}{\omega - \omega' \pm i\eta} \quad (97)$$

from which one obtains

$$\tilde{X}_{nm}^{\pm}(\mathbf{k}, \mathbf{q}; \omega) \equiv \sum_{ij} [M_{nm}^i(\mathbf{k}, \mathbf{q})^*] \tilde{W}_{ij}^{\pm}(\mathbf{q}, \omega) M_{nm}^j(\mathbf{k}, \mathbf{q}) \quad (98)$$

Using the fact that W^c is an even function of ω , it is easy to prove that

$$\tilde{X}_{nm}^{\pm}(\mathbf{k}, \mathbf{q}; -\omega) = -\tilde{X}_{nm}^{\mp}(\mathbf{k}, \mathbf{q}; \omega) \quad (99)$$

We note that it is actually possible to calculate $\tilde{X}_{nm}^{\pm}(\mathbf{k}, \mathbf{q}; \omega)$ directly

$$\tilde{X}_{nm}^{\pm}(\mathbf{k}, \mathbf{q}; \omega) = \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{X_{nm}(\mathbf{k}, \mathbf{q}; \omega')}{\omega - \omega' \pm i\eta} \quad (100)$$

which is likely more efficient since the size of X is usually smaller than W_c so that the number of numerical Hilbert transforms can be reduced. Using \tilde{X}_{nm}^{\pm} , the diagonal correlation self-energy elements at an arbitrary frequency can be calculated as

$$\Sigma_{nk}^c(\omega) = \sum_{\mathbf{q}} \sum_m \tilde{X}_{nm}^{s_m}(\mathbf{k}, \mathbf{q}; \omega - \epsilon_{m\mathbf{k}-\mathbf{q}}) \quad (101)$$

where we use the abbreviation $s_m \equiv \text{sgn}(\epsilon_{m\mathbf{k}-\mathbf{q}} - \epsilon_F)$. In particular, at $\omega = \epsilon_{n\mathbf{k}}$ (the abbreviation $\Delta_{nm} \equiv \epsilon_{n\mathbf{k}} - \epsilon_{m\mathbf{k}-\mathbf{q}}$ is used to simplify the notation)

$$\Sigma_n^c(\epsilon_{n\mathbf{k}}) = \sum_{\mathbf{q}} \sum_m \text{sgn}(\Delta_{nm}) \tilde{X}_{nm}^{s_m \text{sgn}(\Delta_{nm})}(\mathbf{k}, \mathbf{q}; |\Delta_{nm}|) \quad (102)$$

for which Eq. (99) has been used. Since \tilde{X}_{nm}^{\pm} are usually calculated only on a set of discrete points $\{\omega_i\}$, ($i = 1, 2, \dots, N_{\omega}$), to obtain \tilde{X}_{nm}^{\pm} at $\omega = |\Delta_{nm}|$, a linear interpolation can be used; assuming $\omega_i < |\Delta_{nm}| < \omega_{i+1}$, one has

$$\begin{aligned} \tilde{X}^{\pm}(\Delta_{nm}) &= \frac{\omega_{i+1} - |\Delta_{nm}|}{\omega_{i+1} - \omega_i} \tilde{X}^{\pm}(\omega_i) \\ &+ \frac{\omega_i - |\Delta_{nm}|}{\omega_i - \omega_{i+1}} \tilde{X}^{\pm}(\omega_{i+1}) \end{aligned} \quad (103)$$

The derivative of the correlation self-energy at $\omega = \epsilon_{n\mathbf{k}}$ can be calculated as

$$\begin{aligned} \frac{\partial \Re \Sigma_{nk}^c(\epsilon_{n\mathbf{k}})}{\partial \epsilon_{n\mathbf{k}}} &= \sum_{\mathbf{q}} \sum_m \frac{\tilde{X}_{nm}^{s_m}(\mathbf{k}, \mathbf{q}; \omega_{i+1}) - \tilde{X}_{nm}^{s_m}(\mathbf{k}, \mathbf{q}; \omega_i)}{\omega_{i+1} - \omega_i} \end{aligned} \quad (104)$$

3.4.5 The contour deformation approach

Finally we discuss the contour deformation (CD) approach [11,13] to treat the frequency integration for the calculation of the correlation self-energy. The greatest advantage of the CD approach is that the integration is performed along the imaginary frequency, but the self-energy is calculated directly for real frequencies so that the uncertainty related to the analytic continuation can be avoided.

We start from the matrix form of the correlation self-energy, and to simplify the notation, we drop the dependence on wavevectors (\mathbf{k} and \mathbf{q}),

$$\Sigma_n^c(\omega) = \sum_m \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{X_{nm}(\omega')}{\omega + \omega' - \epsilon_m - i\eta \text{sgn}(\epsilon_F - \epsilon_m)} \quad (105)$$

As illustrated in Fig. 2, the integrand in the equation above has two types of poles: 1) the pole from the Green's function, $\omega' = \epsilon_m - \omega + i\eta \text{sgn}(\epsilon_F - \epsilon_m)$, which is above (below) the real frequency for occupied (unoccupied) states; 2) the poles from the screened Coulomb interaction, which are above (below) the real axis for $\omega' < 0$ ($\omega' > 0$). Using the fact $W^c(\omega) \propto \frac{1}{|\omega|^2}$ as $|\omega| \rightarrow \infty$, it is possible to transform the

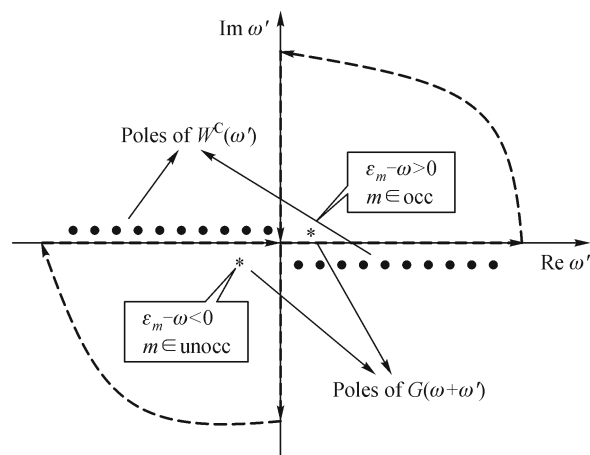


Figure 2 Illustration of the contour deformation approach for the frequency integration

integration over the real frequency to that over the imaginary frequency ($\omega = iu$) plus the contributions from the poles of the Green's function by using the contour deformation technique and the residue theorem [25],

$$\Sigma_n^c(\omega) = \sum_m \left\{ \int_0^\infty \frac{du}{2\pi} X_{nm}(iu) \frac{2(\epsilon_m - \omega)}{(\epsilon_m - \omega)^2 + u^2} + X_{nm}(\epsilon_m - \omega) [\theta(\epsilon_m - \epsilon_F) \theta(\omega - \epsilon_m) - \theta(\epsilon_F - \epsilon_m) \theta(\epsilon_m - \omega)] \right\} \quad (106)$$

In the CD approach, the polarization matrix for both real and imaginary frequencies is needed for the evaluation of the correlation self-energies, so apparently it is more expensive than the IF + AC approach and the Hilbert transform approach. But it is likely the most accurate approach due to the fact that 1) the numerical integration along the imaginary frequency converges very quickly with respect to the number of mesh points ($N_\omega \sim 10 - 20$), in contrast to that along the real frequency for which hundreds of mesh points may be necessary [60]; 2) the correlation self-energy on the real frequency are directly calculated so that it does not rely on the accuracy of the analytic continuation whose error is often difficult to control.

3.5 The choice of the reference H_0

In the spirit of the “best G best W ” strategy, the reference Hamiltonian H_0 should be chosen to deliver the best possible approximation for G and W in the G_0W_0 framework. A certain dependence of the G_0W_0 results on the reference H_0 , although theoretically unsatisfactory, is there unavoidable. When one talks about G_0W_0 , it is important to indicate explicitly the starting point used. For extended sp -electron systems, the LDA/GGA KS single particle Hamiltonian is the most popular choice in practice. We note, however, that using the KS H_0 as the reference for GW calculations is mainly based on pragmatic considerations without much formal justifications since the KS single-particle energies in general cannot be identified as the quasi-particle energies, even at an approximate level. Besides the easy availability of LDA/GGA, there are at least two factors that contribute to the popularity of using the LDA/GGA H_0 : 1) LDA/GGA can often describe the dispersion (with respect to \mathbf{k} -vector) of quasi-particle band structures very well, although the band gaps are significantly underestimated; and 2) The macroscopic dielectric constants of many semiconductors calculated from LDA/GGA eigenenergies and wave functions, which characterizes the strength of screening, are often in good agreement with experiment.

Formally the HF single-particle Hamiltonian is more appropriate as the GW starting point since the HF energies can be related to quasi-particle excitations in terms of Koopmans's theorem [61], although in a quite crude manner. However, the HF descriptions of metals and semiconductors are usually much less satisfactory than the LDA/GGA ones due to its failure to treat the screening (dynamic correlation) effect properly. For wide-gap insulators or finite molecular systems, in which the screening is usually very weak, the HF H_0 might become more appropriate. However, since the implementation of the HF method for extended systems is much more demanding, most G_0W_0 calculations use LDA/GGA H_0 as the starting point.

Besides LDA/GGA Kohn-Sham single-particle Hamiltonian [12,13], other alternative references are also used to improve the starting point for G_0W_0 , including H_0 from the optimized effective potential for exact exchange (OEPx) [62,63], the LDA/GGA plus the Hubbard U correction (LDA/GGA + U) [64–66], and the hybrid functionals approach [67]. The latter two are especially useful for systems with partially occupied d - or f -shells, often termed as strongly correlated systems, for which the LDA/GGA descriptions are often qualitatively wrong so that they are problematic as the reference for G_0W_0 . We refer to Refs. [24,66] for more detailed discussions on this issue.

4 Concluding remarks

In this paper we present an overview of the theoretical foundation of the GW method and main numerical techniques developed for its implementation. The basic theoretical framework for Green's function based approaches has been set up for several decades, but they remain highly specialized and limited to relatively simple systems. The situations are changing quickly in recent years thanks to latest methodological developments as well as the rapid increase of computer power. Even more importantly, there is increasingly stronger demand in fundamental and applied research for accurate theoretical descriptions of electronic band structure. For example, efficient solar energy conversion via photo-catalytic or photo-voltaic processes is currently one of the most intensively pursued scientific frontiers, and electronic band structure is one of the most important parameters of working materials used in solar energy conversion. The Green's function based first-principles methods are therefore highly promising to provide valuable information to interpret existing experimental findings and to achieve rational design of new materials.

There are still several great challenges to meet before it becomes feasible to employ the GF methods routinely for

more complicated systems. The first challenge is related to the efficiency of the current GF methods. In spite of the great efforts that have been invested for the implementation of the *GW* method, including various numerical techniques reviewed in this paper, it is still computationally formidable to routinely treat systems with hundreds of atoms. We expect that some paradigm shift is needed for the implementation of the GF methods that exploits both latest algorithmic advances and new computer architectures. Several interesting developments along this direction have appeared recently [17,19,20,68], but further efforts are obviously needed. We note that in this aspect the experiences from recent developments in linear scaling post-HF quantum chemistry methods [69,70] may give useful inspiration. The second challenge is related to the accuracy of the current GF methods. Even at the *GW* level, additional approximations are involved in practice. With recent proposal of several approximate self-consistent *GW* schemes [14–16], it is still under debate what is the optimal way to perform *GW* calculations. The *GW* is just the lowest order term in the many-body perturbation expansion of the self-energy with respect to *W*. For many systems, in particular, open-shell *d/f*-electron systems, higher order contributions beyond *GW* are probably necessary. In addition, the current GF methods consider only electronic degrees of freedom, and the screening included in *W* has only the contribution from electron-electron interaction, but for some materials, especially those used in solar energy conversion, the electron-phonon coupling can play an important role, and the polaronic contribution to screening may become significant [71–73]. How to incorporate such new important physics in the GF methods is still a frontier not well explored.

Acknowledgements The author acknowledges Dr. Ricardo I. Gomez-Abal, Dr. Patrick Rinke and Prof. Matthias Scheffler for the fruitful collaborations during the stay at Fritz-Haber Institut-der-MPG as a postdoc. The author also thanks Prof. Claudia Ambrosch-Draxel, Prof. E. K. U. Gross, Dr. Xinguo Ren and Dr. Xinzhen Li for helpful discussions. This work was partly supported by NSFC (Project No. 20973009).

References

- Hüfner, S., *Photoelectron Spectroscopy: Principles and Applications* 3rd ed. Berlin: Springer, **2003**
- Onida, G.; Rubio, A., *Rev. Mod. Phys.* **2002**, *74*, 601–659
- Yu, P. Y.; Cardona, M., *Fundamentals of semiconductors: physics and materials properties* 3rd ed. Berlin: Springer, **2001**
- Parr, R. G.; Yang, W., *Density-Functional Theory of Atoms and Molecules* New York: Oxford University Press, **1989**
- Dreizler, R. M.; Gross, E. K. U., *Density Functional Theory: An Approach to the Quantum Many-Body Problem* Berlin: Springer-Verlag, **1990**
- Martin, R. M., *Electronic Structure: Basic Theory and Practical Methods*, Cambridge UK: Cambridge University Press, **2004**
- Aryasetiawan, F., in Anisimov, V. I., ed., *Strong Coulomb Correlations in Electronic Structure Calculations: Beyond the Local Density Approximation* Gordon and Breach Science Publishers, **2000** (1)
- Fetter, A. L.; Walecka, J. D., *Quantum theory of many-particle systems* McGraw-Hill, New York, **1971**
- Hedin, L.; Lundqvist, B. I., *Solid State Phys.* **1969**, *23*, 1–181
- Hedin, L., *Phys. Rev.* **1965**, *139*, A796–A823
- Aryasetiawan, F.; Gunnarsson, O., *Rep. Prog. Phys.* **1998**, *61*, 237–312
- Hybertsen, M. S.; Louie, S. G., *Phys. Rev. B* **1986**, *34*, 5390–5413
- Godby, R. W.; Schlüter, M.; Sham, L. J., *Phys. Rev. B* **1988**, *37*, 10159–10175
- Faleev, S. V.; van Schilfgaarde, M.; Kotani, T., *Phys. Rev. Lett.* **2004**, *93*, 126406
- Bruneval, F.; Vast, N.; Reining, L., *Phys. Rev. B* **2006**, *74*, 045102
- Shishkin, M.; Marsman, M.; Kresse, G., *Phys. Rev. Lett.* **2007**, *99*, 246403
- Bruneval, F.; Gonze, X., *Phys. Rev. B* **2008**, *78*, 085125
- Hamann, D. R.; Vanderbilt, D., *Phys. Rev. B* **2009**, *79*, 045109
- Berger, J. A.; Reining, L.; Sottile, F., *Phys. Rev. B* **2010**, *82*, 2010
- Umari, P.; Stenuit, G.; Baroni, S., *Phys. Rev. B* **2010**, *81*, 115104
- Samsonidze, G.; Jain, M.; Deslippe, J.; Cohen, M. L.; Louie, S. G., *Phys. Rev. Lett.* **2011**, *107*, 186404
- Jiang, H.; Gomez-Abal, R.; Rinke, P.; Scheffler, M., *Phys. Rev. B* **2010**, *81*, 085119
- Inkson, J. C., *Many-body theory of solids: An Introduction* New York: Plenum, **1983**
- Jiang, H. Acta., *Acta. Phys. Chim. Sin* **2010**, *26*, 1017
- Arfken, G. B.; Weber, H. J., *Mathematical Methods for Physicists* ed. 5th ed. Academic Press, **2001**
- Linderberg, J., Öhrn *Propagators in Quantum Chemistry* 2nd ed. John Wiley & Sons, **2004**
- Zakrzewski, V. G.; Dolgounitcheva, O.; Zakjevskii, A. V.; Ortiz, J. V., *Ann. Rep. Comput. Chem* **2010**, *6*, 79–94
- Shishkin, M.; Kresse, G., *Phys. Rev. B* **2007**, *75*, 235102
- Baldereschi, A.; Tosatti, E., *Solid State Commun.* **1979**, *29*, 131–135
- Godby, R. W.; Schlüter, M.; Sham, L. J., *Phys. Rev. B* **1987**, *36*, 6497–6500
- Rojas, H. N.; Godby, R. W.; Needs, R. J., *Phys. Rev. Lett.* **1995**, *74*, 1827–1830
- Rieger, M. M.; Steinbeck, L.; White, I. D.; Rojas, H. N.; Godby, R. W., *Comput. Phys. Commun.* **1999**, *117*, 211–228
- Rohlfing, M.; Krüger, P.; Pollmann, J., *Phys. Rev. Lett.* **1995**, *75*, 3489–3492
- Blase, X.; Attaccalite, C.; Olevano, V. *prb*, **2011**, *83*: 115103
- Helgaker, T.; Jorgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory* John Wiley & Sons, **2000**

36. Foerster, D.; Koval, P.; Sanchez-Portal, D. J., *Chem. Phys.* **2011**, *135*, 074105
37. Gómez-Abal, R.; Li, X.; Scheffler, M.; Ambrosch-Draxl, C., *Phys. Rev. Lett.* **2008**, *101*, 106404
38. Li, X., *All-Electron GOW0 code based on FP-(L)APW+ lo and applications* Ph.D. thesis Free University of Berlin, **2008**
39. Li, G. L.; Yin, Z., *Phys. Chem. Phys. Chem* **2011**, *13*, 2824
40. Aryasetiawan, F., *Phys. Rev. B* **1992**, *46*, 13051–13064
41. Kotani, T.; van Schilfhaarde, M., *Solid State Commun.* **2002**, *121*, 461–465
42. Friedrich, C.; Schindlmayr, A.; Blügel, S.; Kotani, T., *Phys. Rev. B* **2006**, *74*, 045104
43. Friedrich, C.; Blügel, S.; Schindlmayr, A. *prb*, **2010**, *81*: 125102
44. Aryasetiawan, F.; Gunnarsson, O., *Phys. Rev. B* **1994**, *49*, 16214–16222
45. Andersen, O. K., *Phys. Rev. B* **1975**, *12*, 3060–3083
46. Aulbur, W. G.; Jönsson, L.; Wilkins, J. W., *Solid State Phys.* **2000**, *54*, 1–218
47. Gatti, M.; Bruneval, F.; Olevano, V.; Reining, L., *Phys. Rev. Lett.* **2007**, *99*, 266402
48. Vidal, J.; Botti, S.; Olsson, P.; Guillemoles, J.-F.; Reining, L. *prl*, **2010**, *104*: 056401
49. Vidal, J.; Trani, F.; Bruneval, F.; Marques, M. A. L.; Botti, S. *prl*, **2010**, *104*: 136401
50. Botti, S.; Kammerlander, D.; Marques, M. A. L. *apl*, **2011**, *98*: 241915
51. Gygi, F.; Baldereschi, A., *Phys. Rev. Lett.* **1989**, *62*, 2160–2163
52. Massidda, S.; Continenza, A.; Posternak, M.; Baldereschi, A., *Phys. Rev. Lett.* **1995**, *74*, 2323–2326
53. Massidda, S.; Continenza, A.; Posternak, M.; Baldereschi, A., *Phys. Rev. B* **1997**, *55*, 13494–13502
54. Continenza, A.; Massidda, S.; Posternak, M., *Phys. Rev. B* **1999**, *60*, 15699–15704
55. Johnson, D. J., *Phys. Rev. B* **1974**, *9*, 4475–4484
56. Godby, R. W.; Needs, R. J. *prl*, **1989**, *62*: 1169
57. von der Linden, W.; Horsch, P., *Phys. Rev. B* **1988**, *37*, 8351–8362
58. Engel, G. E.; Farid, B., *Phys. Rev. B* **1993**, *47*, 15931–15934
59. Jiang, H.; Engel, E. J., *Chem. Phys.* **2007**, *127*, 184108
60. Shishkin, M.; Kresse, G., *Phys. Rev. B* **2006**, *74*, 035101
61. Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry* New York: McGraw-Hill, **1989**
62. Rinke, P.; Qteish, A.; Neugebauer, J.; Freysoldt, C.; Scheffler, M., *N. J. Phys.* **2005**, *7*, 126
63. Rinke, P.; Qteish, A.; Neugebauer, J.; Scheffler, M. *phys. stat. sol. (b)*, **2008**, *245*: 929
64. Miyake, T.; Zhang, P.; Cohen, M. L.; Louie, S. G., *Phys. Rev. B* **2006**, *74*, 245213
65. Jiang, H.; Gomez-Abal, R. I.; Rinke, P.; Scheffler, M., *Phys. Rev. Lett.* **2009**, *102*, 126403
66. Jiang, H.; Gomez-Abal, R. I.; Rinke, P.; Scheffler, M., *Phys. Rev. B* **2010**, *82*, 045108
67. Rödl, C.; Fuchs, F.; Furthmüller, J.; Bechstedt, F., *Phys. Rev. B* **2008**, *77*, 184408
68. Caramella, L.; Onida, G.; Finocchi, F.; Reining, L.; Sottile, F., *Phys. Rev. B* **2007**, *75*, 205405
69. Schütz, M.; Hetzer, G.; Werner, H. J., *J. Chem. Phys.* **1999**, *111*, 5691
70. Ayala, P. Y.; Scuseria, G. E. J., *Chem. Phys.* **1999**, *110*, 3660
71. Chiodo, L.; Garcia-Lastra, J. M.; Iacomino, A.; Ossicini, S.; Zhao, J.; Petek, H.; Rubio, A. *prb*, **2010**, *82*: 045207
72. Kang, W.; Hybertsen, M. S., *Phys. Rev. B* **2010**, *82*, 085203
73. Wang, H.; Wu, F.; Jiang, H. J., *PhysChemComm* **2011**, *115*, 16180