BERKELEYGW: OVERVIEW & PRACTICAL ISSUES

Felipe H. da Jornada

DEPARTMENT OF MATERIALS SCIENCE AND ENGINEERING STANFORD UNIVERSITY

> DAVID A. STRUBBE DEPARTMENT OF PHYSICS, UC MERCED

Electronic Structure Virtual BerkeleyGW Workshop – June 5th, 2020



Why BerkeleyGW?

- Versatile: supports 3D, 2D, 1D and molecular systems, with Coulomb truncation and efficient k-point sampling algorithms.
- Agnostic: supports a large set of mean-field codes: Quantum ESPRESSO, ABINIT, PARATEC, Octopus, PARSEC, SIESTA, JDFTx, RMGDFT, EPM.
- **General**: support for **semiconductor**, **metallic** and **semi-metallic** systems.
- **Massively parallel**: scales to **512,000 CPU cores**, supports distributed memory and hybrid architectures. High-performance GPU support to be released. Can handle large systems containing thousands of atoms.
- Free & open source.





Del Ben, Jornada, Deslippe, Louie, CPC 235, 187 (2018).





Over 2,000 downloads since July 2018





Substrate renormalization of the self-energy and excitonic effects



Qiu, Jornada, Louie, Nano Lett., 17, 4706 (2017).

Defects in monolayer TMDs



Long-lived dispersionless plasmons in quasi-2D metals



Jornada, Xian, Rubio, Louie, Nat. Comm 11, 1013 (2020).





BerkeleyGW philosophy

> What we strive for:

- Code correctness
- Supporting a diversity of mean-field codes (Quantum ESPRESSO, Abinit, Paratec, JDFTx, RMGDFT, PARSETC etc.)
- Implementing scalable codes & algorithms
- Supporting new physics / features

> What we do not focus on:

Black-box design



1. Introduction

- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation
- 7. Real and complex versions



1. Introduction

- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation
- 7. Real and complex versions



1. Introduction: GW calculations – Theory

Electronic self-energy Σ within the GW approximation:

2

Screened Coulomb interaction W:

RPA dielectric matrix:

Noninteracting polarizability matrix:

 $\Sigma(\mathbf{r},\mathbf{r}';t) \approx iG_0(\mathbf{r},\mathbf{r}';t)W_0(\mathbf{r},\mathbf{r}';t)$

 $W_{\mathbf{G}\mathbf{G}'}^{0}(\mathbf{q};\omega) \equiv \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q};\omega)v(\mathbf{q}+\mathbf{G}')$

Sum over wavevectors **q**!

$$\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G}\mathbf{G}'} - \nu(\mathbf{q}+\mathbf{G})\chi^{0}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$$

$$\chi^0(\mathbf{r},\mathbf{r}';t) = G_0(\mathbf{r},\mathbf{r}';t)G_0(\mathbf{r}',\mathbf{r};-t)$$

Noninteracting Green's function (spectral representation)

$$G_0(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{n\mathbf{k}} \frac{\phi_{n\mathbf{k}}(\mathbf{r})\phi_{n\mathbf{k}}^*(\mathbf{r}')}{\omega - E_{n\mathbf{k}}^{\mathrm{mf}} \pm i\eta}$$

Sum over all occupied and unoccupied states *n*!



1. Introduction: GW calculations – Practice





1. Introduction

2. K-point vs. Q-point grids

- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation
- 7. Real and complex versions



2. k-grids and q-grids

$$\chi^{0}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega=0) \sim \sum_{v \ c \ \mathbf{k}} \frac{M^{*}_{v\mathbf{k}+\mathbf{q},c\mathbf{k}}(\mathbf{G}) \ M_{v\mathbf{k}+\mathbf{q},c\mathbf{k}}(\mathbf{G}')}{E_{v\mathbf{k}+\mathbf{q}} - E_{c\mathbf{k}}}$$

$$M_{\nu\mathbf{k}+\mathbf{q},c\mathbf{k}}(\mathbf{G}) \equiv \langle v\mathbf{k} + \mathbf{q} | e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | c\mathbf{k} \rangle$$

Mean-field quantities: computed in any regular
 <u>k-point grid</u> (does not need to be Γ centered)



Polarizability & dielectric matrices: computed in a regular, Γ-centered <u>q-point grid</u>, with q = k' - k



In practice: often use the same **F-centered grid** for both k and q points (esp. bulk systems)

2. **k**-grids and **q**-grids: **q**=0 point

$$\blacktriangleright$$
 Take **G** = **G**' = 0 and ω = 0:

$$\varepsilon_{00}(\mathbf{q},0) = 1 - \frac{4\pi e^2}{|\mathbf{q}|^2} \chi_{00}^0(\mathbf{q},0)$$

> Cannot directly compute
$$\varepsilon(\mathbf{q} = \mathbf{0})!$$

- For gapped systems:
 - > Compute $\varepsilon(\mathbf{q}_0)$ at a small but finite $\mathbf{q}_0 \sim 0.001$.
 - Use two sets of k-point grids and wave functions for valence/conduction states.
- > WFN: provides conduction states for $\varepsilon(\mathbf{q} = \mathbf{q}_0)$ + all states for $\mathbf{q} \neq \mathbf{q}_0$.

> WFNq: provides valence states for $\varepsilon(\mathbf{q} = \mathbf{q}_0)$.



q₀-shifted **k**'-point grid (WFNq)





2. **k**-grids and **q**-grids



2. Specification of q-points in epsilon.inp



- ✤ Metals
 - Screening depends critically on sampling DOS at Fermi surface for intraband transitions. Cannot use shifted grid! See manual / 2019 BerkeleyGW Workshop

Sigma, Kernel, Absorption use q-grid defined by eps0mat.h5 / epsmat.h5



2. Example k-grid construction: 4 × 4 grid for graphene

How to construct k-point grid (WFN) and q-shifted k-point grids (WFNq)?

> kgrid.x utility!

Example on the right:

- o **Graphene**
- 4x4x1 Monkhorst-Pack grid
- $\mathbf{q}_0 = (0.0, 0.05, 0.0)$

 \mathbf{q}_0 shift breaks symmetry and gives more points.

Main grid (<mark>WFN</mark>) 16 in full BZ Reduced to 4

Shifted grid (WFNq) 16 in full BZ Reduced to 10





shifted



2. k-, q-grids and bands

For reference: simplified approach for tutorial

	k-grid	# bands	Comments
SCF	uniform, no shift	occupied	
WFN	uniform, no shift	many	
WFNq	WFN + q -shift	occupied	
epsilon.inp q -points	WFN but q ₀ instead of 0	many	bands to sum over
WFN_inner	WFN	many	bands to sum over
sigma.inp k -points	subset of WFN_inner	few	can choose to calculate Sigma just for bands of interest
WFN_co	WFN_inner	few	
WFN_fi (absorption)	uniform, random shift	few	
WFNq_fi	WFN_fi + q -shift	occupied	
WFN_fi (inteqp)	anything	few	whatever is of interest

- 1. Introduction
- 2. K-point vs. Q-point grids
- 3. Screening models for $\epsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation
- 7. Real and complex versions



3. Divergent behavior of
$$W_{\mathbf{G},\mathbf{G}'}$$
 for $\mathbf{q} = 0$
 $v(\mathbf{q})$ DOS inter/intra-band transitions
 $\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega=0) \sim \frac{1}{|\mathbf{q}+\mathbf{G}|^2} \sum_{v \in \mathbf{k}} \frac{M_{vk+q,ck}^*(\mathbf{G}) M_{vk+q,ck}(\mathbf{G}')}{E_{vk+q} - E_{ck}}$
 $W_{\mathbf{G},\mathbf{G}'}(\mathbf{q};\omega) \sim \frac{1}{|\mathbf{q}+\mathbf{G}|^2} \sum_{v \in \mathbf{k}} \frac{M_{vk+q,ck}^*(\mathbf{G}) M_{vk+q,ck}(\mathbf{G}')}{E_{vk+q} - E_{ck}}$
 $W_{\mathbf{G},\mathbf{G}'}(\mathbf{q};\omega) = \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q};\omega) v(\mathbf{q}+\mathbf{G})$
 $W_{\mathbf{G},\mathbf{G}'}(\mathbf{q};0) \text{ for } \mathbf{q} \rightarrow 0:$
 \cdot Diverges for semiconductors
 \cdot Is finite for metals
Note: integral of $W_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ over \mathbf{q} is finite
 q^2/q^2
See BerkeleyGW paper: arXiv:1111.4429.
Head: $\mathbf{G}=0, \mathbf{G}'=0$
 $W_{\mathrm{ing}}: \mathbf{G}=0, \mathbf{G}'=0$
 $\mathbf{W}_{\mathrm{ing}}: \mathbf{G}=0, \mathbf{G}'$

3. Solution: screening models

Use calculation of $\varepsilon(\mathbf{q}_0 \sim 0.001)$ along the periodic direction to **parametrize screening model** for $\mathbf{q} \approx 0$.

The calculation is still *ab initio*! The screening model is just a "hint" the the user give to BerkeleyGW to improve w.r.t. k-point sampling!

Input file epsilon.inp:

screening_semiconductor
#screening_graphene
#screening_metal

See BerkeleyGW paper <u>arXiv:1111.4429</u> and <u>manual</u>.

Sigma: Use model to perform Monte Carlo integration $\int d^3 q M(\mathbf{q}) W_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ over region around $\mathbf{q} = 0$

Absorption: interpolate kernel

$$\left\langle vc\mathbf{k} \left| K \right| v'c'\mathbf{k}' \right\rangle = \frac{a_{vc\mathbf{k}v'c'\mathbf{k}'}}{A\left(\mathbf{q}\right)} + \frac{b_{vc\mathbf{k}v'c'\mathbf{k}'}}{B\left(\mathbf{q}\right)} + \frac{c_{vc\mathbf{k}v'c'\mathbf{k}'}}{C\left(\mathbf{q}\right)}$$

Note: anisotropic materials need to use direction such that $\epsilon^{-1}(\mathbf{q}_0) = \langle \epsilon^{-1}(\mathbf{q}) \rangle$



3. Coulomb truncation: different screening models

0D (e.g.: molecule) fully confined



cell_box_truncation

1D (e.g.: nanotube) periodic along z



2D (e.g.: graphene) periodic along x,y

 $v_{\rm t} =$



cell_slab_truncation

See BerkeleyGW paper <u>arXiv:1111.4429</u> and <u>manual</u>.



- 1. Introduction
- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation
- 7. Real and complex versions



4. Frequency dependence of $\varepsilon(\omega)$

Computationally expensive part of the calculation: correlation contribution to Σ:

$$\Sigma^{\text{Cor}}(\mathbf{r},\mathbf{r}';\omega) = -\frac{1}{\pi} \sum_{n} \psi_n(\mathbf{x}) \psi_n^*(\mathbf{x}') \int_0^\infty d\omega' W_{\text{aH}}(\mathbf{x},\mathbf{x}';\omega') \times \frac{1}{\omega - \epsilon_n - (\omega' - i\eta) \operatorname{sgn}(\epsilon_n - \epsilon_F)}$$

- Computationally expensive: need $\varepsilon(\omega)$ for all frequencies
- Extra convergence parameters

Simplification: plasmon-pole model (PPM)

- Compute $\varepsilon(\omega = 0)$, use physical models & constraints to obtain $\varepsilon(\omega \neq 0)$: charge density **RHO**.
- PPM is the default option (and a good idea for a first calculation).

```
"full-frequency" vs. "plasmon-pole"
```

Input file epsilon.inp:

#frequency_dependence 0

See BerkeleyGW paper <u>arXiv:1111.4429</u> and <u>manual</u>.



- 1. Introduction
- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$

5. Symmetry and degeneracy

- 6. Solving Dyson's equation
- 7. Real and complex versions



5. Degeneracy check utility

Choice of bands can break symmetry of degenerate states and lead to arbitrary results:

ERROR: Selected number of bands breaks degenerate subspace.

- Use degeneracy_check.x utility to find number of bands that does not break degeneracies.
- So, could use **number_bands 32** in Epsilon.
- Can also turn off degeneracy enforcement (degeneracy_check_override flag).
 Ok if include many unoccupied bands (error from breaking deg. subspace vanishes)

\$ degeneracy_check.x WFN

8 14

18

20 32

Number of spins:

Number of bands:

Number of k-points:

Reading eigenvalues from file WFN

1

8

== Degeneracy-allowed numbers of bands (for epsilon and sigma) ==

Note: cannot assess whether or not highest band 35 is degenerate.

35

1. Introduction

- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy

6. Solving Dyson's equation

7. Real and complex versions



6. Solving Dyson's equation in Sigma

$$E_{n\mathbf{k}}^{\mathrm{QP}} = E_{n\mathbf{k}}^{\mathrm{MF}} + \langle \psi_{n\mathbf{k}} | \Sigma(E_{n\mathbf{k}}^{\mathrm{QP}}) - \Sigma^{\mathrm{MF}} | \psi_{n\mathbf{k}} \rangle$$

How can we solve when we don't know *E*^{QP} yet?

(1) eqp0: evaluate at
$$E^{MF}$$
. $E_{n\mathbf{k}}^{QP0} = E_{n\mathbf{k}}^{MF} + \langle \psi_{n\mathbf{k}} | \Sigma(E_{n\mathbf{k}}^{MF}) - \Sigma^{MF} | \psi_{n\mathbf{k}} \rangle$

(2) eqp1: solve linearized approximation (Newton's Method)

$$E_{n\mathbf{k}}^{\text{QP1}} = E_{n\mathbf{k}}^{\text{QP0}} + \frac{d\Sigma_{n\mathbf{k}}/dE}{1 - d\Sigma_{n\mathbf{k}}/dE} \left(E_{n\mathbf{k}}^{\text{QP0}} - E_{n\mathbf{k}}^{\text{MF}} \right)$$

Available as columns in **sigma_hp.log**, and **eqp0.dat** and **eqp1.dat** files

X Note: for full frequency calculations, **eqp1** reports the full numerical solution of Dyson's equation.

6. Mean-field exchange-correlation functional

$$E_{n\mathbf{k}}^{\mathrm{QP}} = E_{n\mathbf{k}}^{\mathrm{MF}} + \langle \psi_{n\mathbf{k}} | \Sigma(E_{n\mathbf{k}}^{\mathrm{QP}}) - \Sigma^{\mathrm{MF}} | \psi_{n\mathbf{k}} \rangle$$

If we start from KS DFT, $\Sigma - \Sigma^{MF} = \Sigma^{GW} - V_{XC}$

BerkeleyGW accepts two forms of V_{XC} :

- **VXC**: Binary file containing the operator in G space: $V_{XC}(\mathbf{r}) \rightarrow V_{XC}(\mathbf{G})$
 - BerkeleyGW can compute arbitrary matrix elements given the operator.
- **vxc.dat**: ASCII file containing the matrix elements in a KS orbitals: $\langle n\mathbf{k}|V_{XC}|n'\mathbf{k}\rangle$.
 - DFT code must compute all matrix elements for the relevant states.

SerkeleyGW also natively supports some hybrid functionals even if the DFT wrapper cannot write $\langle n\mathbf{k}|\Sigma_x|n'\mathbf{k}\rangle$. See flag **bare_exchange_fraction** for the Sigma code.

1. Introduction

- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation

7. Real and complex versions



7. Real or complex flavor?

e.g. epsilon.real.x, epsilon.cplx.x

Complex is general, but real is faster, uses less memory and disk space

Real: only with inversion symmetry about the origin $u(-\mathbf{r}) = au(\mathbf{r})$ and time-reversal symmetry $u^*(\mathbf{r}) = bu(\mathbf{r})$ a, b each equal to ± 1

Plane-wave expansion:

$$u(\mathbf{r}) = \sum_{\mathbf{G}} u_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \longrightarrow u_{\mathbf{G}} = cu_{\mathbf{G}}^* \qquad \succ \text{ Can choose } c = 1 \text{ for real coefficients}$$

Same for density and $V_{\rm xc}$, except no need for time-reversal $\rho(\mathbf{r}) = \rho^*(\mathbf{r})$.

- What breaks time-reversal? Fractional translations, magnetic fields, spin-polarization, spinors.
- Plane-wave codes generally just use complex wavefunctions.
- Conditions for reality depends on the basis! Real-space: *k* = 0, time-reversal.



Questions?

- 1. Introduction
- 2. K-point vs. Q-point grids
- 3. Screening models for $\varepsilon(\mathbf{q})$
- 4. Frequency dependence of $\varepsilon(\omega)$
- 5. Symmetry and degeneracy
- 6. Solving Dyson's equation
- 7. Real and complex versions

Homework: try installing code on your own cluster!

Extra slides: discussion of half-shifted grids, calculations for metals



k-, q-grids and bands

recommended approach (using half-shifted grids)

	k-grid	# bands	Comments
SCF	Uniform, 0.5 shift	occupied	as usual in DFT
WFN	Uniform, 0.5 shift	many	
WFNq	WFN + q -shift	occupied	
epsilon.inp q -points	WFN but no shift, q ₀	many	bands to sum over
WFN_inner	WFN but no shift	many	bands to sum over
sigma.inp k -points	subset of WFN_inner	few	can choose to calculate Sigma just for bands of interest
WFN_co	WFN_inner	few	
WFN_fi (absorption)	Uniform, random shift	few	
WFNq_fi	WFN_fi + q -shift	occupied	
WFN_fi (inteqp)	anything	few	whatever is of interest

5. Choice of bands can break symmetry in GW/BSE

Symmetry operations are only defined for non-degenerate states!

Example:

- 2D system with mirror symmetry $\hat{\sigma}_x$ about the x=0 plane.
- Consider two degenerate KS states of $|p_x\rangle$ and $|p_y\rangle$ characters:

 $\circ \ \hat{\sigma}_{x} | p_{x} \rangle = - | p_{x} \rangle$ $\circ \ \hat{\sigma}_{x} | p_{y} \rangle = | p_{y} \rangle$

- If $|p_x\rangle$ and $|p_y\rangle$ degenerate, the DFT code generates arbitrary linear combination, e.g. $|p_x\rangle \pm i |p_y\rangle$, which are not an eigenstate of $\hat{\sigma}_x$.
- > Including only some of a degenerate space will break symmetry.
- Results depends on arbitrary linear combinations in mean-field. Not reproducible!



k-grid construction: 4 × 4 shifted grid for graphene



k-grid construction: 4 × 4 shifted grid for graphene





Quasiparticle renormalization factor Z



A. Damascelli, Z. Hussain, Z.-X Shen, Rev. Mod. Phys. 75, 473 (2003)

Special treatment for metals

Coefficients depend critically on sampling DOS at Fermi surface for intraband transitions.

Two separate runs of Epsilon

eps0mat: $\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}\left(\mathbf{q}_{0}
ight)$

WFN = WFNq only a small number of bands for intra-band transitions around Fermi surface very fine: grid spacing is $\mathbf{q}_0 e.g.$ grid = 32 × 32 × 32 unshifted, $\mathbf{q}_0 = (0, 0, 1/32)$

epsmat:
$$\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}\left(\mathbf{q}\neq\mathbf{q}_{0}
ight)$$

$\epsilon_{\mathbf{GG}'}^{-1}$	head	wing	wing'	body
Semiconductor	const	q	\mathbf{q}/q^2	const
Metal	q^2	q^2	const	const
$W_{{f G}{f G}'}$	head	wing	wing'	body
				v
Semiconductor	$1/q^{2}$	\mathbf{q}/q^2	\mathbf{q}/q^2	const

WFN = WFNq. unshifted, many bands, ordinary fineness. *e.g.* grid = 12 × 12 × 12.



jornada@stanford.edu • https://jornada.stanford.edu







epsmat: $\epsilon_{\mathbf{G}\mathbf{G}'}^{-1} (\mathbf{q} \neq \mathbf{q}_0)$



•••

end

Problem 1: Non-smooth behavior around q = 0

(14, 0) carbon nanotube wire truncation

Systems with reduced dimensionality are harder to converge!

Not covered in this tutorial See example 4 from 2019 BerkeleyGW Workshop



