The GW Code of ABINIT

Present status, new features and future developments

M. Giantomassi[1]

[1]Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Liége, 31 Jan 2007

1. Introduction
   - The Kohn Sham Band-Gap Problem
   - Hedin's Equations and the GW Approximation
2. GW Implementation
   - Formalism in Frequency and Reciprocal Space
   - Implementation
   - Symmetries: a Useful Tool to Speed Up the Code
3. Future Developments

1. Introduction
   - The Kohn Sham Band-Gap Problem
   - Hedin's Equations and the GW Approximation
2. GW Implementation
   - Formalism in Frequency and Reciprocal Space
   - Implementation
   - Symmetries: a Useful Tool to Speed Up the Code
3. Future Developments

1. Introduction
   - The Kohn Sham Band-Gap Problem
   - Hedin's Equations and the GW Approximation
2. GW Implementation
   - Formalism in Frequency and Reciprocal Space
   - Implementation
   - Symmetries: a Useful Tool to Speed Up the Code
3. Future Developments

The Kohn-Sham (KS) approach to DFT is powerful and successful because:

- KS equations are easy to solve
- E_{xc} is suitable for approximations

BUT...

(1) KS eigenvalues and eigenfunctions have no direct physical meaning

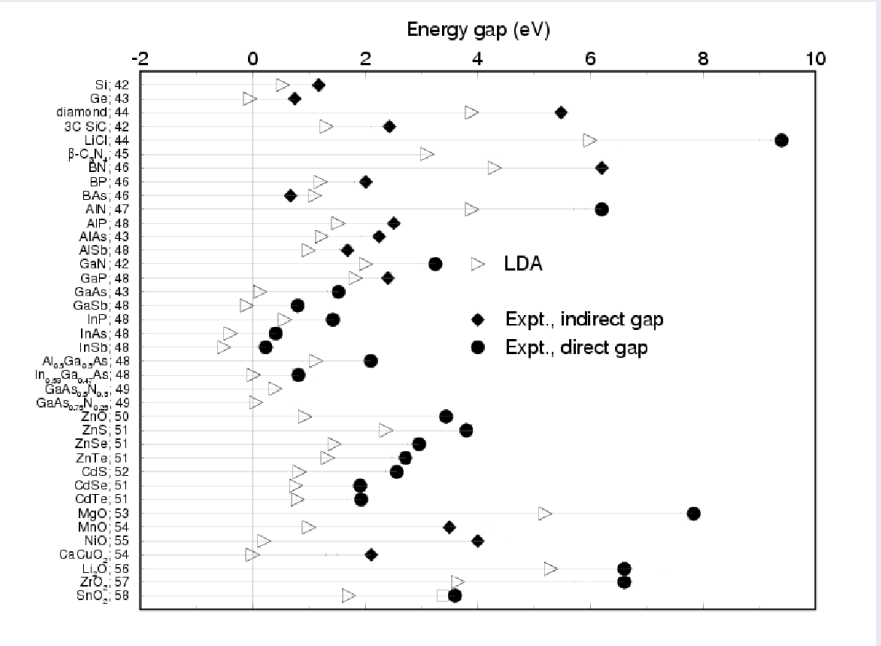
(2) KS energies cannot be interpreted as addition/removal energies

(3) Exact DFT gives the ionization potential, but approximated $\hat{v}_{xc}$'s give poor results

Nevertheless they are used to study band structures!

It turns out that:

- KS yelds reasonable valence band dispersions but. . .
- KS sistematically underestimates the band gap value!



Direct and indirect LDA gaps compared to exp. values. After W. G. Aulbur *et al.*

Better theoretically justified approaches are mandatory to study band gaps!

What about Many Body Theory?

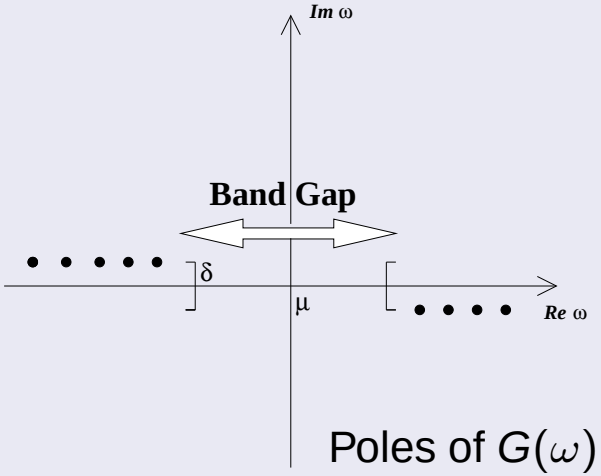
Green's Function

$$G(1,2) := -i\langle N|\hat{T}\hat{\psi}_{\mathcal{H}}(1)\hat{\psi}^{\dagger}_{\mathcal{H}}(2)|N\rangle$$

no spin-flip, no spin-orbit coupling

$$G^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2;\omega) = \sum_{\alpha} \frac{\Psi^{\sigma}_{\alpha}(\mathbf{r}_1)\Psi^{\sigma\,\dagger}_{\alpha}(\mathbf{r}_2)}{\omega - E^{\sigma}_{\alpha} + i\delta\,\mathrm{sign}(E^{\sigma}_{\alpha} - \mu)}$$

- Poles of $G(\omega)$ are addition or removal energies (QP energies, E^{σ}_{α})
- No poles inside the energy gap
- Ψ^{σ}_{α} are the quasiparticle amplitudes satisfying the Dyson equation:



Poles of $G(\omega)$

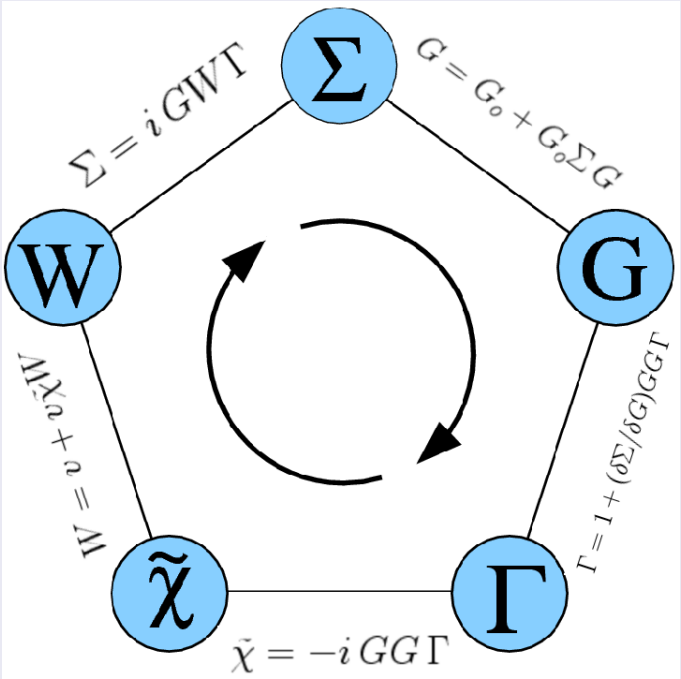
$$\left[-\frac{1}{2}\Delta + \hat{v}_{\mathrm{H}} + \hat{v}_{\mathrm{ext}}\right]\Psi^{\sigma}_{\alpha}(\mathbf{r}_1) + \int \Sigma^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2;E^{\sigma}_{\alpha})\Psi^{\sigma}_{\alpha}(\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_2 = E^{\sigma}_{\alpha}\Psi^{\sigma}_{\alpha}(\mathbf{r}_1)$$

$\Sigma^{\sigma\sigma}$ is a non-Hermitian, nonlocal and energy dependent operator (Self-Energy)

Σ can be obtained by solving a closed set of five integral equations involving:

- The Green's functions G
- The screened interaction $W = \epsilon^{-1} v$
- The irreducible polarizability $\tilde{\chi} = \dfrac{\delta n}{\delta v_{\text{tot}}}$
- The vertex function Γ

Graphical representation of Hedin's equations



- Challenging task!
- Find an approximation that can be solved
- Idea: neglect vertex corrections

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

- Approximated Self Energy:

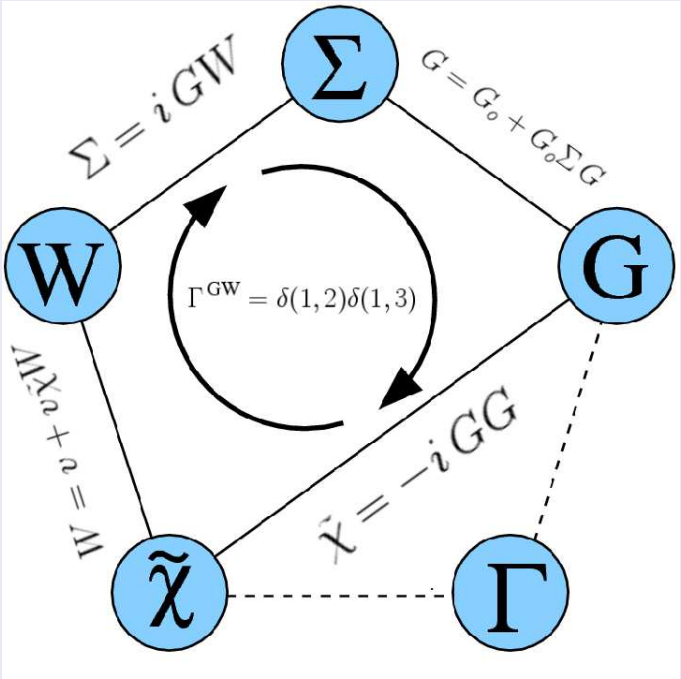
$$\Sigma = iGW$$

Hence GW Approximation!

Σ can be obtained by solving a closed set of five integral equations involving:

- The Green's functions G
- The screened interaction $W = \epsilon^{-1} v$
- The irreducible polarizability $\tilde{\chi} = \dfrac{\delta n}{\delta v_{\text{tot}}}$
- The vertex function Γ

The GW Approximation



- Simpler set of equations:
- Start from a good G_o and iterate
- Different level of self-consistency:
  - (A) G_oW_o (only one cycle)
  - (B) G_oW (update only W)
  - (C) GW_o (update only G)
  - (D) GW (fully self-consistent)

GW Self-Energy in ω-space:

$$\Sigma^{\sigma\sigma}_{\mathrm{GW}}(\mathbf{r}_1,\mathbf{r}_2;\omega) = \frac{i}{2\pi}\int e^{i\omega'\delta} G^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2;\omega+\omega') W(\mathbf{r}_1,\mathbf{r}_2;\omega')\,\mathrm{d}\omega'$$

First Iteration (G_oW_o)

$G^{\sigma\sigma}$ is approximated by the independent particle $G_o^{\sigma\sigma}$
(the Green's function of the non-interacting KS system)

$$G_o^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2;\omega) = \sum_\alpha \frac{\psi^\sigma_\alpha(\mathbf{r}_1)\psi^{\sigma\,\dagger}_\alpha(\mathbf{r}_2)}{\omega-\epsilon^\sigma_\alpha + i\delta\,\mathrm{sign}(\epsilon^\sigma_\alpha-\mu)}$$

with:

$$\hat{H}_{\mathrm{KS}}\,\psi^\sigma_\alpha = \epsilon^\sigma_\alpha\psi^\sigma_\alpha.$$

Ingredients: KS wavefunctions and eigenvalues

GW Self-Energy in ω-space:

$$\Sigma^{\sigma\sigma}_{\mathrm{GW}}(\mathbf{r}_1,\mathbf{r}_2;\omega) = \frac{i}{2\pi}\int e^{i\omega'\delta} G^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2;\omega+\omega') W(\mathbf{r}_1,\mathbf{r}_2;\omega')\,\mathrm{d}\omega'$$

First Iteration (G_oW_o)

W is approximated by the RPA expression:

$$W_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega) = \epsilon^{-1}_{\mathbf{G}_1\mathbf{G}_2} v_{\mathbf{G}_2}(\mathbf{q}), \quad \text{where} \quad \epsilon^{\mathrm{RPA}}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega) = \delta_{\mathbf{G}_1\mathbf{G}_2} - v_{\mathbf{G}_1}(\mathbf{q})\chi^{(0)}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega)$$

and

$$\chi^{(0)}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega) = \frac{1}{V}\sum_{\substack{\mathbf{k}\sigma\\ b_1 b_2}} \frac{\left[f(\epsilon^{\sigma}_{\mathbf{k}-\mathbf{q}b_2}) - f(\epsilon^{\sigma}_{\mathbf{k}b_1})\right] M^{b_2 b_1}_{\mathbf{G}_1}(\mathbf{q},\mathbf{k},\sigma) {M^{b_2 b_1}_{\mathbf{G}_2}}^{\dagger}(\mathbf{q},\mathbf{k},\sigma)}{\omega + \epsilon^{\sigma}_{\mathbf{k}-\mathbf{q}b_2} - \epsilon^{\sigma}_{\mathbf{k}b_1} - i\delta\,\mathrm{sign}(\epsilon^{\sigma}_{\mathbf{k}-\mathbf{q}b_2} - \epsilon^{\sigma}_{\mathbf{k}b_1})}$$

$$M^{\nu c}_{\mathbf{G}}(\mathbf{q},\mathbf{k},\sigma) := \langle \mathbf{k}-\mathbf{q},\nu,\sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \mathbf{k},c,\sigma\rangle$$

Ingredients: KS wavefunctions, eigenvalues and occupations numbers f

Schematic representation of a RPA calculation

(1) FFT to evaluate $M^{\nu c}_{\mathbf{G}}(\mathbf{q},\mathbf{k},\sigma)$

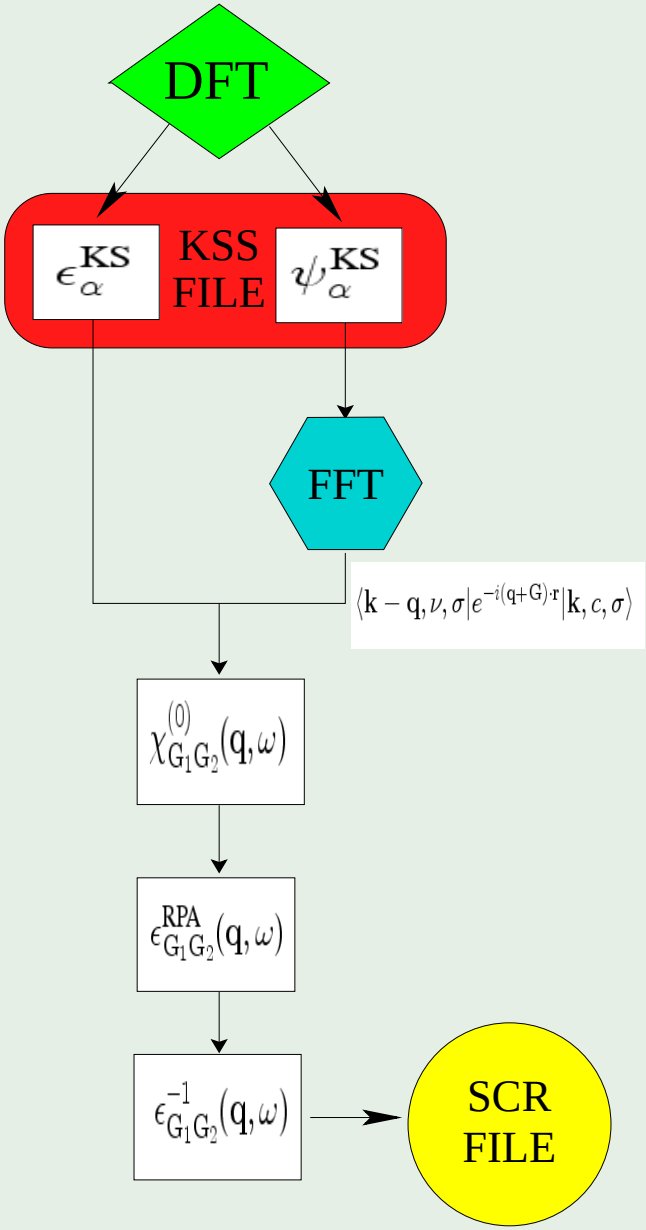
- Efficient
- Can use less PW's in $|b_1 \mathbf{k}\,\sigma\rangle$

(2) Evaluation of $\chi^{(0)}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega)$ scales as $N_\sigma \cdot N_\nu \cdot N_c \cdot N_\omega \cdot N_{\mathbf{k}}$

- Time consuming ($N_c \sim 10^2$)
- Parallelized over **k**'s in the BZ
- Parallelized over N_c (v5.3.0)

(3) Inversion of ϵ^{RPA}

- Done by a sigle CPU
- Could be parallelized



Oscillator matrix elements

$$M_{\mathbf{G}}^{\nu c}(\mathbf{q},\mathbf{k},\sigma) := \langle \mathbf{k}-\mathbf{q},\nu,\sigma|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\mathbf{k},c,\sigma\rangle$$

Matrix elements of $\Sigma_x^{\sigma\sigma}$ (From the static part of W)

$$\langle b_1\mathbf{k}\,\sigma|\Sigma_x^{\sigma\sigma}|b_2\mathbf{k}\,\sigma\rangle = -\frac{4\pi}{V}\sum_{\nu}^{\text{occ}}\sum_{\mathbf{q}\,\mathbf{G}}\frac{M_{\mathbf{G}}^{\nu b_1\,\dagger}(\mathbf{k},\mathbf{q},\sigma)M_{\mathbf{G}}^{\nu b_2}(\mathbf{k},\mathbf{q},\sigma)}{|\mathbf{q}+\mathbf{G}|^2}$$

Matrix elements of $\Sigma_c^{\sigma\sigma}(\omega)$ (From the frequency dependent part of W)

$$\langle b_1\mathbf{k}\sigma|\Sigma_c^{\sigma\sigma}(\omega)|b_2\mathbf{k}\sigma\rangle = \sum_{\substack{\mathbf{q}\\ \mathbf{G}_1\mathbf{G}_2}}\sum_{n}\frac{M_{\mathbf{G}_1}^{nb_1\,\dagger}(\mathbf{q},\mathbf{k},\sigma)M_{\mathbf{G}_2}^{nb_2}(\mathbf{q},\mathbf{k},\sigma)}{|\mathbf{q}+\mathbf{G}_1||\mathbf{q}+\mathbf{G}_2|}\cdot J_{\mathbf{G}_1\mathbf{G}_2}^{n,\mathbf{k}-\mathbf{q}}(\mathbf{q},\omega,\sigma)$$

Schematic representation of a Σ calculation

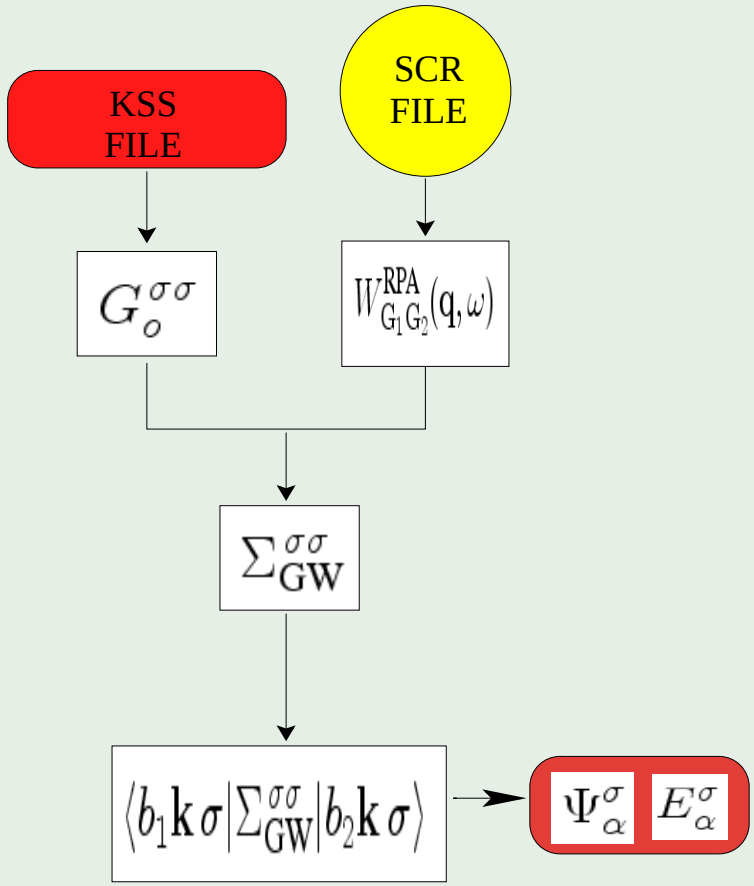
(1) Only Σ matrix elements are calculated

- Save memory and CPU
- But it could be useful to have Σ and/or G!

(2) $\Sigma_x^{\sigma\sigma}$ scales as $N_\nu \cdot N_{\mathbf{G}} \cdot N_{\mathbf{q}}$

$\Sigma_c^{\sigma\sigma}$ scales as $(N_\nu + N_c) \cdot N_{\mathbf{G}}^2 \cdot N_{\mathbf{q}}$

- Parallelized over **q** in the BZ
- Parallelized over $(N_\nu + N_c)$ (v5.3.0)



Parallelization works well, but we are still summing over the full BZ

Symmetries can be used to speed up the code. . .

Definitions:

- A symmetry operation $\hat{S}_{\mathbf{t}}$ consists of:
  (1) a proper or improper rotation S
  (2) a fractional translation $\mathbf{t}$
- We adopt the convention: $\hat{S}\ \psi(\mathbf{r}) := \psi(S^{-1}(\mathbf{r}))$
- S is an element of the little group $L_{\mathbf{q}}$ if $S\mathbf{q} = \mathbf{q}$
- $\mathrm{IBZ}_{\mathbf{q}}$ is the irreducible wedge defined by $L_{\mathbf{q}}$
- If $\mathbf{t} = \overrightarrow{0}$ the symmetry operation is called symmorphic

Only symmorphic operations will be treated

Symmetry properties:

If n is not degenerate

$$\psi_{nS\mathbf{k}}(\mathbf{r}) = \psi_{n\mathbf{k}}(S^{-1}\mathbf{r})$$

If n belongs to the degenerate subspace $\mathcal{C}_n$

$$\psi_{nS\mathbf{k}}(\mathbf{r}) = \sum_{\alpha}^{\mathcal{C}_n} D_{n\alpha}(S)\,\psi_{\alpha\mathbf{k}}(S^{-1}\mathbf{r})$$

where $D_{n\alpha}(S)$ is the unitary transformation associated with S

Using these properties in the definition of $M^{\nu c}_{\mathbf{G}}(\mathbf{q},\mathbf{k})$, we obtain:

If $S\mathbf{q} = \mathbf{q}$, ν and c are not degenerate

$$M^{\nu c}_{\mathbf{G}}(\mathbf{q}, S\mathbf{k}) = M^{\nu c}_{S^{-1}\mathbf{G}}(\mathbf{q},\mathbf{k})$$

If $S\mathbf{q} = \mathbf{q}$ and only ν is degenerate

$$M^{\nu c}_{\mathbf{G}}(\mathbf{q}, S\mathbf{k}) = \sum_{\alpha}^{\mathcal{C}_\nu} D^{\dagger}_{\nu\alpha}(S)\, M^{\alpha c}_{S^{-1}\mathbf{G}}(\mathbf{q},\mathbf{k})$$

Symmetrization in $\chi^{(0)}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega)$

- The external wavevector $\mathbf{q}$ defines the irreducible wedge for the integration

Schematically:

$$\chi^{(0)}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega) \propto \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{\substack{\sigma \\ b_1 b_2}} F_{\mathbf{G}_1\mathbf{G}_2} \Longrightarrow \sum_{\mathbf{k}}^{\mathrm{IBZ}_\mathbf{q}} w(\mathbf{k}) \sum_{\substack{\sigma \\ b_1 b_2}} \tilde{F}_{\mathbf{G}_1\mathbf{G}_2}$$

where:

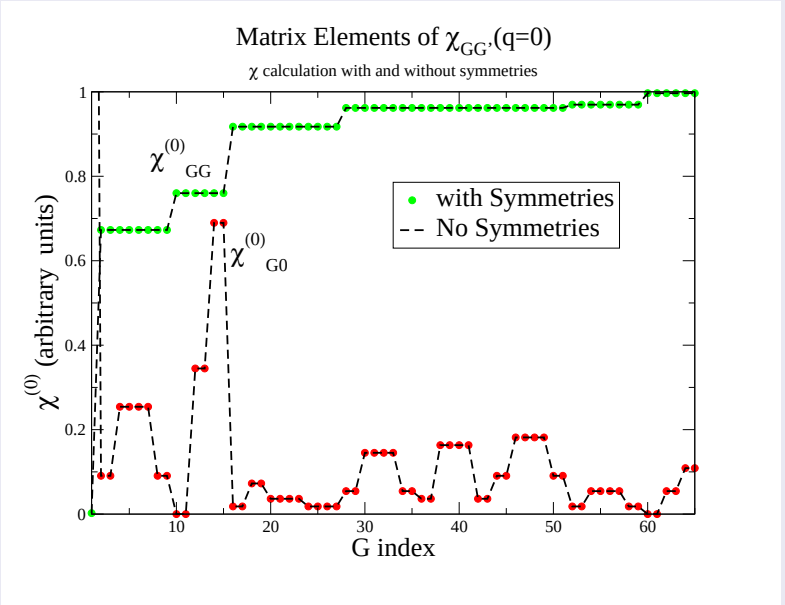
$w(\mathbf{k})$ are appropriate weights for each point in $\mathrm{IBZ}_\mathbf{q}$

$\tilde{F}_{\mathbf{G}_1\mathbf{G}_2}$ is the symmetrized expression corresponding to $F_{\mathbf{G}_1\mathbf{G}_2}$

- The use of symmetries reduces the effort for high-symmetric $\mathbf{q}$-points
- No need to take into account the transformation $D_{\alpha\beta}(S)$ (matrices $D_{\alpha\beta}(S)$ drop out of the final expression)

RPA calculation for fcc Ni
(48 symmetry operations)

q	$N_{\mathbf{k}}$ Sym/No Sym	CPU Time Sym/No Sym
$\longrightarrow \Gamma$	29/512	30m/9.6h
X	120/512	9m/31m
$\frac{1}{4}, \frac{3}{8}, \frac{3}{8}$	288/512	18m/31m



On a single processor a χ calculation with symmetries requires 8.5h instead of 15.2h

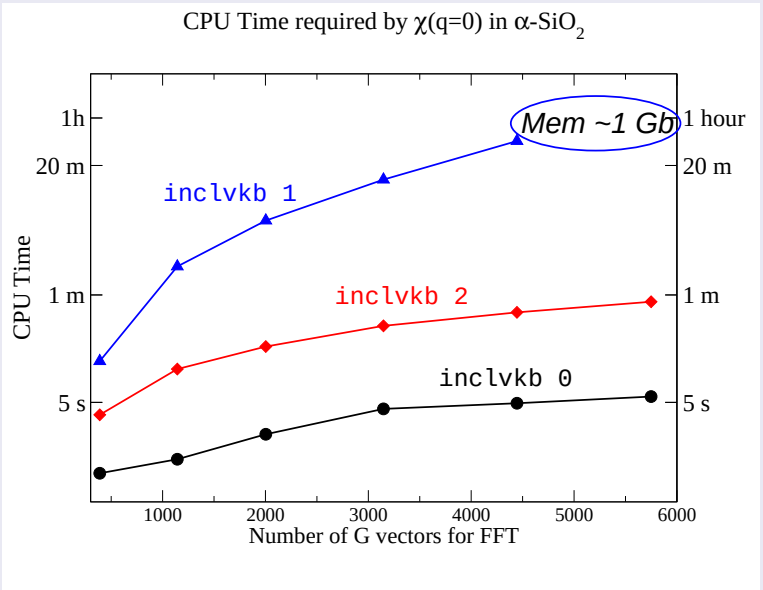
- Symmetrization decreases the CPU time but not the memory
- Why is $\mathbf{q} \longrightarrow \Gamma$ the most time consuming point?

Correct evaluation of $\lim_{\mathbf{q}\to 0} \chi^{(0)}_{\mathbf{G}0}(\mathbf{q},\omega)$ and $\lim_{\mathbf{q}\to 0} \chi^{(0)}_{00}(\mathbf{q},\omega)$ requires:

$$V_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{k}) := \left(\nabla_{\mathbf{k}+\mathbf{G}_1} + \nabla_{\mathbf{k}+\mathbf{G}_2}\right) \langle \mathbf{k}+\mathbf{G}_1 | \hat{V}_{\text{nl}} | \mathbf{k}+\mathbf{G}_2 \rangle$$

Three approaches according to inclvkb:

- $0 \Longrightarrow$ do not include $V_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{k})$
- $1 \Longrightarrow$ use Legendre polymomials:
  - (i) $V_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{k})$ cannot be factorized
  - (ii) CPU and memory $\propto N_{\mathbf{G}}^2$
- $2 \Longrightarrow$ use spherical harmonics:
  - (i) $V_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{k}) = \tilde{V}_{\mathbf{G}_1}(\mathbf{k})\tilde{V}_{\mathbf{G}_2}(\mathbf{k})$
  - (ii) CPU and memory $\propto \gamma \cdot N_{\mathbf{G}}$



Symmetrization in Σ

- The external wavevector **k** defines the irreducible wedge for the integration

Schematically:

$$\langle b_1 \mathbf{k}\,\sigma|\Sigma^{\sigma\sigma}|b_1 \mathbf{k}\,\sigma\rangle \propto \sum_{\mathbf{q}}^{\mathrm{BZ}} \sum_{\substack{n \\ \mathbf{G}_1\mathbf{G}_2}} F \Longrightarrow \sum_{\mathbf{q}}^{\mathrm{IBZ}_{\mathbf{k}}} w(\mathbf{q}) \sum_{\substack{n \\ \mathbf{G}_1\mathbf{G}_2}} \tilde{F}$$

BUT…

- If b_1 is degenerate then $D_{b_1\alpha}(S)$ should be included in the equations
- A naive symmetrization causes removals of degeneracies

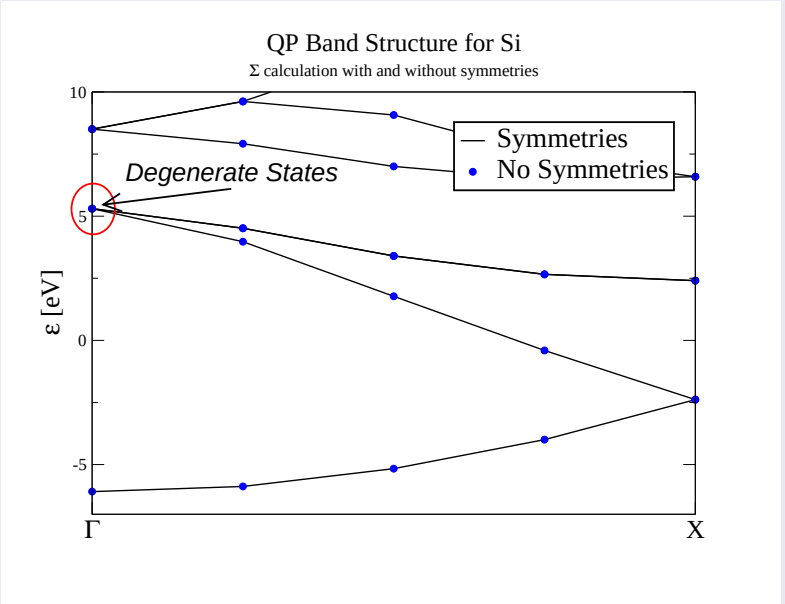
Solution:

Include all the degenerate states in the Σ calculation and average QP energies

But what about the off-diagonal elements of Σ?

QP band structure calculation for Si (24 symmetry operations)

k	$N_{\mathbf{q}}$ Sym/No Sym	CPU Time Sym/No Sym
Γ	19/256	8m/1.5h
X	60/256	21m/1.5h
$\frac{1}{2}, \frac{3}{8}, \frac{1}{16}$	160/256	0.9h/1.5h



On a single processor a Σ calculation in 9 k-points requires 4.5h instead of 14h

- Symmetrization decreases the CPU time, but not the memory allocated
- Small drawback: degenerate states must be included in the Σ calculation

Future Developments

(1) Inclusion of *Umklapp processes*: $S\mathbf{q} = \mathbf{q} + \mathbf{G}$

- further reduction of **k**- and **q**-points in the sums

(2) Treatment of non-symmorphic operations to:

- reduce memory
- improve the symmetrization

(3) Symmetries + Different levels of parallelization:

- k-points and spin
- bands and spin

(4) Better alghoritms to compute $\chi^{(0)}_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q},\omega)$