Calculation of response properties in DFT

Gaspard Kemlin

gaspard.kemlin@enpc.fr
CERMICS, ENPC & Inria Paris, team MATHERIALS

PhD student with É. Cancès & A. Levitt,
joint work with M. F. Herbst and B. Stamm

GAMM jahrestagung, Wednesday 17th 2022, Aachen

Why computing response to external perturbations ?

- Kohn–Sham Density Functional Theory (KS-DFT) ⇝ directly gives quantities of interest such as ground-state density and energy.

- However, many quantities of interest depends on the response of the system to external perturbations:
  - forces (response to atomic displacements) are easy thanks to the Hellmann-Feynman theorem;
  - in general, one needs to compute the response of the orbitals to external perturbations (phonons, polarisability, conductivity, . . .);
  - machine learning applications require derivative w.r.t. model parameters.

DFT and response calculations ○●○
Computation of response ○○○○○○
Numerical tests ○○○○○○○○
Take-home messages and outlooks ○○

DFT and response calculations

The Kohn–Sham equations for a system with $N_{el} = 2N_p$ electrons read

$$
H_\rho \phi_n = \varepsilon_n \phi_n, \quad \varepsilon_1 \leqslant \varepsilon_2 \leqslant \cdots
$$

$$
\langle \phi_n, \phi_m \rangle = \delta_{nm}
$$

$$
\rho(\boldsymbol{r}) = \sum_{n=1}^{+\infty} f_n \left| \phi_n(\boldsymbol{r}) \right|^2, \quad \sum_{n=1}^{+\infty} f_n = N_{el}
$$

where

- $H_\rho = -\frac{1}{2}\Delta + V + V_H(\rho) + V_{xc}(\rho)$ is the Kohn–Sham Hamiltonian;
- $f_n \in [0, 2]$ is the occupation number of the orbital ϕ_n:
  - for insulators and semi-conductors, $f_n = \begin{cases} 2 & \text{if } n \leqslant N_p, \\ 0 & \text{if } n > N_p; \end{cases}$
  - for metals, one usually uses finite temperature T and $f_n = f\left(\dfrac{\varepsilon_n - \varepsilon_F}{T}\right)$, with f a fixed *smearing* function (e.g. $f(x) = 2/(1 + e^x)$). ε_F is then defined such that $\displaystyle\sum_{n=1}^{+\infty} f_n = N_{el}$.

DFT and response calculations
○○●

Computation of response
○○○○○○

Numerical tests
○○○○○○○○

Take-home messages and outlooks
○○

Assume that you have computed a solution to the Kohn–Sham equations. How does the density ρ changes if the Hamiltonian is perturbed by an external potential δV ?

<hr>

[1]S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. Reviews of Modern Physics, 73(2):515–562, 2001.

[2]M. F. Herbst and A. Levitt. Black-box inhomogeneous preconditioning for self-consistent field iterations in density functional theory. Journal of Physics: Condensed Matter, 33(8):085503, 2020.

[3]A. Levitt. Screening in the Finite-Temperature Reduced Hartree–Fock Model. Archive for Rational Mechanics and Analysis, 238(2):901–927, 2020.

DFT and response calculations
○○●

Computation of response
○○○○○○

Numerical tests
○○○○○○○○

Take-home messages and outlooks
○○

Assume that you have computed a solution to the Kohn–Sham equations. How does the density ρ changes if the Hamiltonian is perturbed by an external potential δV ?

In this framework, the response to an external perturbation δV can be computed[123] *via*

$$\delta\rho(\boldsymbol{r}) = \sum_{n=1}^{+\infty} \sum_{m=1}^{+\infty} \frac{f_n - f_m}{\varepsilon_n - \varepsilon_m} \phi_n^*(\boldsymbol{r})\phi_m(\boldsymbol{r}) \left(\delta V_{mn} - \delta\varepsilon_F \delta_{mn}\right),$$

where $A_{mn} = \langle\phi_m, A\phi_n\rangle$. We use the convention

$$\frac{f_n - f_n}{\varepsilon_n - \varepsilon_n} = \frac{1}{T} f' \left(\frac{\varepsilon_n - \varepsilon_F}{T}\right) =: f_n'.$$

---

[1]S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. Reviews of Modern Physics, 73(2):515–562, 2001.

[2]M. F. Herbst and A. Levitt. Black-box inhomogeneous preconditioning for self-consistent field iterations in density functional theory. Journal of Physics: Condensed Matter, 33(8):085503, 2020.

[3]A. Levitt. Screening in the Finite-Temperature Reduced Hartree–Fock Model. Archive for Rational Mechanics and Analysis, 238(2):901–927, 2020.
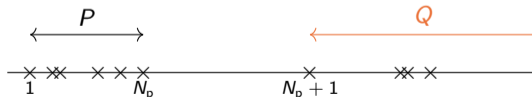
Insulators and semi-conductors

For insulators and semi-conductors, things are easy:

$$\delta\rho(\boldsymbol{r}) = 2\sum_{n=1}^{N_\mathrm{p}}\sum_{m=N_\mathrm{p}+1}^{+\infty}\frac{2}{\varepsilon_n - \varepsilon_m}\phi_n^*(\boldsymbol{r})\phi_m(\boldsymbol{r})\delta V_{mn} = 2\sum_{n=1}^{N_\mathrm{p}}\phi_n(\boldsymbol{r})^*\delta\phi_n(\boldsymbol{r}),$$

where $\delta\phi_n(\boldsymbol{r})$ can be computed from the Sternheimer equation[4]

$$Q(H_\rho - \varepsilon_n)Q\delta\phi_n = -2Q\delta V\phi_n,$$

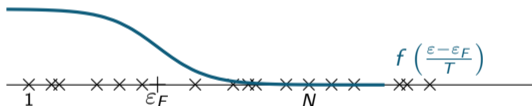where $Q = 1 - P$ and $P = \sum_{m=1}^{N_\mathrm{p}}|\phi_m\rangle\langle\phi_m|$.

---

[4]R. M. Sternheimer. Electronic Polarizabilities of Ions from the Hartree-Fock Wave Functions. Physical Review, 96(4):951–968, 1954.

DFT and response calculations
000

Computation of response
0●0000

Numerical tests
00000000

Take-home messages and outlooks
00

Metals

The real fun happens with metals:

- first, select N orbitals that have an occupation number f_n above some numerical threshold;

DFT and response calculations
○○○

Computation of response
○●○○○○○

Numerical tests
○○○○○○○○○

Take-home messages and outlooks
○○

Metals

The real fun happens with metals:

- first, select N orbitals that have an occupation number f_n above some numerical threshold;
- introduce free parameters $\alpha_{mn} \in [0, 1]$ such that $\alpha_{mn} + \alpha_{nm} = 1$. Using symmetry between n and m, we obtain

$$\delta\rho(\boldsymbol{r}) = 2\sum_{n=1}^{N}\sum_{m=1}^{+\infty}\frac{f_n - f_m}{\varepsilon_n - \varepsilon_m}\alpha_{mn}\phi_n^*(\boldsymbol{r})\phi_m(\boldsymbol{r})\left(\delta V_{mn} - \delta\varepsilon_F\delta_{mn}\right) = 2\sum_{n=1}^{N}\phi_n^*(\boldsymbol{r})\delta\phi_n(\boldsymbol{r});$$

Metals

The real fun happens with metals:

- first, select N orbitals that have an occupation number f_n above some numerical threshold;
- introduce free parameters $\alpha_{mn} \in [0, 1]$ such that $\alpha_{mn} + \alpha_{nm} = 1$. Using symmetry between n and m, we obtain

$$\delta\rho(\boldsymbol{r}) = 2 \sum_{n=1}^{N} \sum_{m=1}^{+\infty} \frac{f_n - f_m}{\varepsilon_n - \varepsilon_m} \alpha_{mn} \phi_n^*(\boldsymbol{r}) \phi_m(\boldsymbol{r}) \left(\delta V_{mn} - \delta\varepsilon_F \delta_{mn} \right) = 2 \sum_{n=1}^{N} \phi_n^*(\boldsymbol{r}) \delta\phi_n(\boldsymbol{r});$$

- conservation of charge directly gives $\int \delta\rho(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} = 0 \Rightarrow \delta\varepsilon_F = \left(\sum_{n=1}^{N} f_n' \delta V_{nn} \right) / \left(\sum_{n=1}^{N} f_n' \right)$;

Metals

The real fun happens with metals:

- first, select N orbitals that have an occupation number f_n above some numerical threshold;
- introduce free parameters $\alpha_{mn} \in [0, 1]$ such that $\alpha_{mn} + \alpha_{nm} = 1$. Using symmetry between n and m, we obtain

$$\delta\rho(\boldsymbol{r}) = 2\sum_{n=1}^{N}\sum_{m=1}^{+\infty}\frac{f_n - f_m}{\varepsilon_n - \varepsilon_m}\alpha_{mn}\phi_n^*(\boldsymbol{r})\phi_m(\boldsymbol{r})\left(\delta V_{mn} - \delta\varepsilon_F\delta_{mn}\right) = 2\sum_{n=1}^{N}\phi_n^*(\boldsymbol{r})\delta\phi_n(\boldsymbol{r});$$

- conservation of charge directly gives $\int \delta\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = 0 \Rightarrow \delta\varepsilon_F = \left(\sum_{n=1}^{N} f_n' \delta V_{nn}\right) \Big/ \left(\sum_{n=1}^{N} f_n'\right)$;
- for $n \leqslant N$, split $\delta\phi_n$ into two contributions: $\delta\phi_n = \delta\phi_n^P + \delta\phi_n^Q$ where

  - $\delta\phi_n^P = \displaystyle\sum_{m=1}^{N}\langle\phi_m, \delta\phi_n\rangle\,\phi_m \in \mathsf{Span}(\phi_m)_{1\leqslant m\leqslant N}$ can be explicitly computed;
  - $\delta\phi_n^Q = \displaystyle\sum_{m=N+1}^{+\infty}\langle\phi_m, \delta\phi_n\rangle\,\phi_m \in \mathsf{Span}(\phi_m)_{N+1\leqslant m}$ can be obtained through the Sternheimer equation.

$\delta\phi_n^P = \displaystyle\sum_{m=1}^{N} \langle \phi_m, \delta\phi_n \rangle \, \phi_m$ can be obtained by computing all the contributions

$$\langle \phi_m, \delta\phi_n \rangle = \frac{f_n - f_m}{\varepsilon_n - \varepsilon_m} \alpha_{mn} \left(\delta V_{mn} - \delta\varepsilon_F \delta_{mn} \right).$$

Different possibilities for α_{mn} exist (because the Sternheimer equation is ill-posed in $\mathrm{Span}(\phi_m)_{1 \leqslant m \leqslant N}$):

- $\alpha_{mn} = 1/2$ is the simplest possibility;
- $\alpha_{mn} = f_n^2/(f_n^2 + f_m^2)$ makes $\delta\phi_n$ small if f_n is small (implemented for instance in DFTK);
- $\alpha_{mn} = f\left(\frac{\varepsilon_n - \varepsilon_m}{T} \right)$ where $f(x) = 1/(1 + \mathrm{e}^x)$ (implemented in Quantum Espresso);
- $\alpha_{mn} = \mathbf{1}_{f_n > f_m}$ and $\alpha_{nn} = 1/2$ (implemented in Abinit);
- whatever you like as long as $\alpha_{mn} + \alpha_{nm} = 1$.

DFT and response calculations

000

Computation of response

000●00

Numerical tests

00000000

Take-home messages and outlooks

00

$\delta\phi_n^P = \sum_{m=1}^{N} \langle \phi_m, \delta\phi_n \rangle \phi_m$ can be obtained by computing all the contributions

$$\langle \phi_m, \delta\phi_n \rangle = \frac{f_n - f_m}{\varepsilon_n - \varepsilon_m} \alpha_{mn} \left(\delta V_{mn} - \delta\varepsilon_F \delta_{mn} \right).$$

Different possibilities for α_{mn} exist (because the Sternheimer equation is ill-posed in $\mathrm{Span}(\phi_m)_{1 \leqslant m \leqslant N}$):

- $\alpha_{mn} = 1/2$ is the simplest possibility;
- $\alpha_{mn} = f_n^2/(f_n^2 + f_m^2)$ makes $\delta\phi_n$ small if f_n is small (implemented for instance in DFTK);
- $\alpha_{mn} = f\left(\frac{\varepsilon_n - \varepsilon_m}{T}\right)$ where $f(x) = 1/(1 + \mathrm{e}^x)$ (implemented in Quantum Espresso);
- $\alpha_{mn} = \mathbf{1}_{f_n > f_m}$ and $\alpha_{nn} = 1/2$ (implemented in Abinit);
- whatever you like as long as $\alpha_{mn} + \alpha_{nm} = 1$.

This is nice because it brings numerical stability:

$$\left| \langle \phi_m, \delta\phi_n \rangle \right| \leqslant \max_{x \in \mathbb{R}} \frac{1}{T} \left| f'(x) \right| \left| \delta V_{mn} - \delta\varepsilon_F \delta_{mn} \right|,$$

so that an error on δV is amplified at most by $\max_{x \in \mathbb{R}} \frac{1}{T} |f'(x)|$.

$\delta\phi_n^Q = \displaystyle\sum_{m=N+1}^{+\infty} \langle \phi_m, \delta\phi_n \rangle \, \phi_m$ cannot be computed in a similar way as we do not know all the ϕ_m for $m \geqslant N + 1$. However, as for insulators, it solves the Sternheimer equation:

$$Q(H_\rho - \varepsilon_n)Q\delta\phi_n = -f_n Q\delta V \phi_n,$$

where $Q = 1 - \displaystyle\sum_{m=1}^{N} |\phi_m\rangle \langle \phi_m|$.

---

[5]S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. Reviews of Modern Physics, 73(2):515–562, 2001.

$\delta\phi_n^Q = \displaystyle\sum_{m=N+1}^{+\infty} \langle\phi_m, \delta\phi_n\rangle \phi_m$ cannot be computed in a similar way as we do not know all the ϕ_m for $m \geqslant N+1$. However, as for insulators, it solves the Sternheimer equation:

$$Q(H_\rho - \varepsilon_n)Q\delta\phi_n = -f_n Q\delta V\phi_n,$$

where $Q = 1 - \displaystyle\sum_{m=1}^{N} |\phi_m\rangle\langle\phi_m|$.

⤳ This can be solved with iterative solvers, but it is possibly very ill-conditioned as, for metals, $\varepsilon_{N+1} - \varepsilon_N > 0$ can be very small. Solutions to this exists in the literature (e.g. appropriate shift of the Hamiltonian[5]), and we suggest a new one, based on the inversion of a (better conditioned) Schur complement.

---

[5]S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. Reviews of Modern Physics, 73(2):515–562, 2001.

DFT and response calculations
ooo

Computation of response
oooooeo

Numerical tests
oooooooo

Take-home messages and outlooks
oo

We actually have some information about additional orbitals $\widetilde{\Phi} = (\widetilde{\phi}_m)_{N+1 \leqslant m \leqslant N+N_{ex}}$:

- some of them have been discarded from the response calculations because f_n is too small, these are exact (up to numerical tolerance) eigenvectors;
- others have been used to enhance the convergence of the SCF algorithm, but they have not been fully converged by the eigensolver.

We actually have some information about additional orbitals $\widetilde{\Phi} = (\widetilde{\phi}_m)_{N+1 \leqslant m \leqslant N+N_{\mathrm{ex}}}$:

- some of them have been discarded from the response calculations because f_n is too small, these are exact (up to numerical tolerance) eigenvectors;
- others have been used to enhance the convergence of the SCF algorithm, but they have not been fully converged by the eigensolver.

In particular, we can assume that $\langle \widetilde{\Phi}, H_\rho \widetilde{\Phi} \rangle$ is a diagonal matrix. We can thus write H_ρ into the decomposition
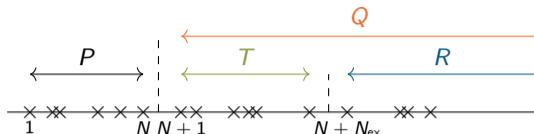
$$\mathrm{Ran}(P) \oplus \mathrm{Ran}(T) \oplus \mathrm{Ran}(R) \quad \text{with} \quad P = \sum_{m=1}^{N} |\phi_m\rangle \langle \phi_m|, \ \ T = \sum_{m=N+1}^{N+N_{\mathrm{ex}}} |\widetilde{\phi}_m\rangle \langle \widetilde{\phi}_m|, \ \ R = 1 - P - T.$$

as

$$H_\rho = \begin{pmatrix} E & 0 & 0 \\ 0 & E_{\mathrm{ex}} & TH_\rho R \\ 0 & RH_\rho T & RH_\rho R \end{pmatrix}$$

where

$$E = \mathrm{Diag}(\varepsilon_1, \ldots, \varepsilon_N), \quad E_{\mathrm{ex}} = \langle \widetilde{\Phi}, H_\rho \widetilde{\Phi} \rangle, \quad RH_\rho T = 0 \text{ if } \widetilde{\Phi} \text{ is exclusively exact eigenvectors.}$$

$$\mathrm{Ran}(P) \oplus \mathrm{Ran}(Q)$$

$$H_\rho - \varepsilon_n = \begin{pmatrix} E - \varepsilon_n & 0 \\ 0 & Q(H_\rho - \varepsilon_n)Q \end{pmatrix}$$

invert $Q(H_\rho - \varepsilon_n)Q$

$\downarrow$

possibly ill-conditioned for $n = N$
if $\varepsilon_{N+1} - \varepsilon_N$ is too small

$$\mathrm{Ran}(P) \oplus \mathrm{Ran}(T) \oplus \mathrm{Ran}(R)$$

$$H_\rho - \varepsilon_n = \begin{pmatrix} E - \varepsilon_n & 0 & 0 \\ 0 & E_{\mathrm{ex}} - \varepsilon_n & T(H_\rho - \varepsilon_n)R \\ 0 & R(H_\rho - \varepsilon_n)T & R(H_\rho - \varepsilon_n)R \end{pmatrix}$$

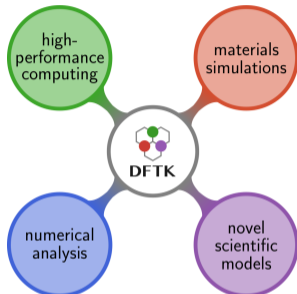invert $E_{\mathrm{ex}} - \varepsilon_n$ for free (diagonal)

invert $R(H_\rho - \varepsilon_n)R$

$\downarrow$

better conditioned for $n = N$ as $\varepsilon_{N+N_{\mathrm{ex}}+1} - \varepsilon_N > \varepsilon_{N+1} - \varepsilon_N$

plug things together *via* a Schur complement to get $\delta\phi_n^Q$

DFT and response calculations
○○○
Computation of response
○○○○○○○
Numerical tests
●○○○○○○○○
Take-home messages and outlooks
○○

Density-functional toolkit[6] – `https://dftk.org`

- `Julia` code for plane-wave DFT
- Fully composable with `Julia` ecosystem:
    - Arbitrary precision
    - Automatic Differentiation
    - Numerical error control

- Both suitable for mathematical developments and relevant applications
    - 1D problems, toy models for rigorous analysis
    - DFT up to $1,000$ electrons

- 3 years of development (M.F. Herbst and A. Levitt) and $\sim$ 7k lines of code

---

[6] M. F. Herbst, A. Levitt, and E. Cancès. DFTK: A Julian approach for simulating electrons in solids. Proceedings of the JuliaCon Conferences, 3(26):69, 2021.

Aluminium

Al_{40}: elongated aluminium supercell with 40 atoms and we use

- PBE exchange-correlation functional;
- Fermi-Dirac smearing with $T = 10^{-3}$ hartree;
- $3 \times 3 \times 1$ discretization of the Brillouin zone;
- $E_{cut} = 45$ hartree;
- $N_p = 60$ electron pairs $\rightsquigarrow$ standard heuristics give 72 bands + 3 nonconverged bands for every k-point for calculations, occupation threshold is 10^{-8}.

$\rightsquigarrow$ we compute $\delta\phi_n^P$ with $\alpha_{mn} = f_n^2/(f_n^2 + f_m^2)$, then solve the Sternheimer equation for $\delta\phi_n^Q$ with and without the Schur complement, for every Bloch fiber of the periodic KS Hamiltonian associated to each k-point.

k-point		1		2		5	
N		69		58		67	
N_{ex}		6		17		8	
ε_{N-1}	f_{N-1}	0.359	$1.03 \cdot 10^{-5}$	0.359	$7.77 \cdot 10^{-6}$	0.344	1.85
ε_N	f_N	0.359	$1.02 \cdot 10^{-5}$	0.360	$5.23 \cdot 10^{-6}$	0.344	1.84
ε_{N+1}	f_{N+1}	0.391	$1.25 \cdot 10^{-19}$	0.373	$8.01 \cdot 10^{-12}$	0.366	$9.16 \cdot 10^{-9}$
$\varepsilon_{N+1} - \varepsilon_N$		0.0320		0.0134		0.0217	
#iterations $n = N$ Schur		42		**41**		37	
#iterations $n = N$ no Schur		49		**74**		53	

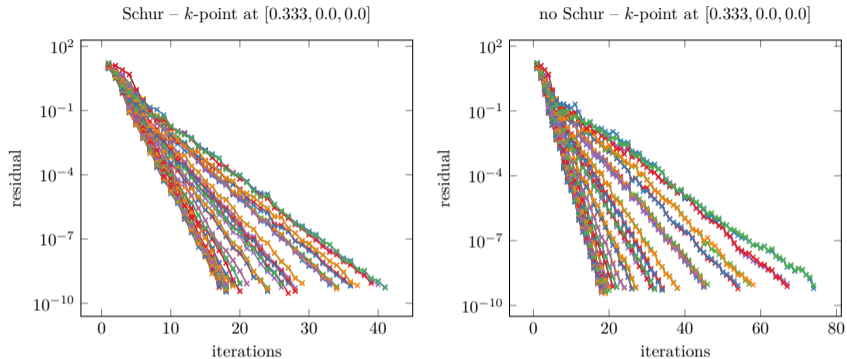Table: Sternheimer convergence data for 3 particular k-points for Al_{40}.

Figure: Sternheimer convergence data for all eigenvalues of one particular k-point for Al_{40}.

⤳ global computational time (all k-points included) is reduced from $8,090$ applications of the Hamiltonian without the Schur complement to $6,960$ (15% gain).

DFT and response calculations
000

Computation of response
000000

Numerical tests
00000●000

Take-home messages and outlooks
00

Heusler compounds

Fe$_2$MnAl: halfmetallic behavior, spin $\uparrow \sim$ metal and spin $\downarrow \sim$ insulator

- PBE exchange-correlation functional;
- Gaussian smearing with $T = 10^{-2}$ hartree;
- $13 \times 13 \times 13$ discretization of the Brillouin zone;
- $E_{\text{cut}} = 45$ hartree;
- $f_n \in [0, 1]$ here, but we double the number of k-point (one for each spin);
- $N_{\text{p}} = 25$ electron pairs $\rightsquigarrow$ standard heuristics give 35 bands + 3 nonconverged bands for every k-point for calculations, occupation threshold is 10^{-8}

$\rightsquigarrow$ we compute $\delta\phi_n^P$ with $\alpha_{mn} = f_n^2/(f_n^2 + f_m^2)$, then solve the Sternheimer equation for $\delta\phi_n^Q$ with and without the Schur complement, for every Bloch fiber of the periodic KS Hamiltonian associated to each k-point.

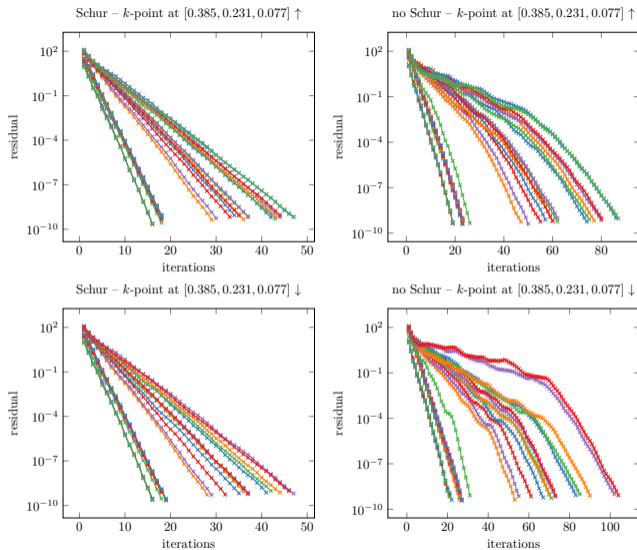| spin channel		↑		↓	
N		28		26	
N_{ex}		10		12	
ε_{N-2}	f_{N-2}	0.447	0.877	0.438	0.992
ε_{N-1}	f_{N-1}	0.469	0.0213	0.480	0.00016
ε_N	f_N	0.473	0.00608	0.491	$1.72 \cdot 10^{-7}$
ε_{N+1}	f_{N+1}	0.515	$1.06 \cdot 10^{-17}$	0.506	$1.8 \cdot 10^{-13}$
$\varepsilon_{N+1} - \varepsilon_N$		0.0423		0.0154	
#iterations $n = N$ Schur		**47**		**47**	
#iterations $n = N$ no Schur		**87**		**104**	

Table: Sternheimer convergence data for both spin channel of one particular k-points for Fe_2MnAl.

Figure: Sternheimer convergence data for all eigenvalues of both spin channels of one particular k-point for Fe₂MnAl.
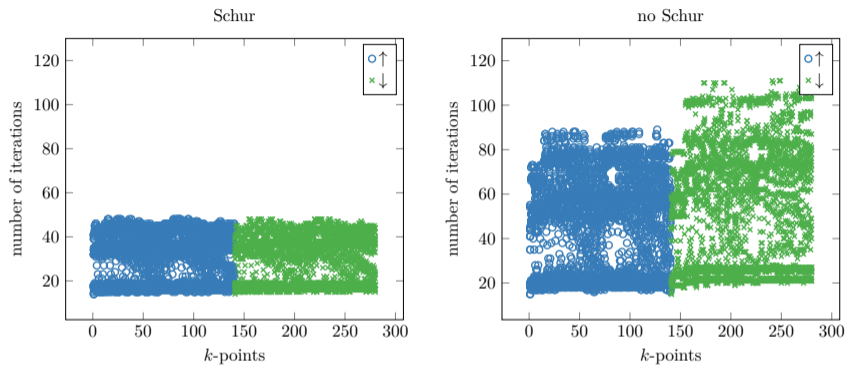
Schur | no Schur

Figure: Sternheimer convergence data for all eigenvalues of both spin channels of all k-points for Fe_2MnAl.

⤳ global computational time (all k-points included) is reduced from 83.7×10^3 applications of the Hamiltonian without the Schur complement to 56.1×10^3 (33% gain).

Take-home messages and outlooks

Take-home messages:

- insulators are easy: $\delta\phi_n \in \text{Span}(\phi_m)_{N+1\leqslant m}$ and the Sternheimer equation is usually well-conditioned;
- metals are more difficult: $\delta\phi_n = \delta\phi_n^P + \delta\phi_n^Q$
  - $\delta\phi_n^P$ solves the ill-posed Sternheimer equation in $\text{Span}(\phi_m)_{1\leqslant m\leqslant N}$ and we derived a common framework from the literature which ensures numerical stability (computational time is negligible);
  - $\delta\phi_n^Q$ solves the ill-conditioned Sternheimer equation in $\text{Span}(\phi_m)_{N+1\leqslant m}$ and we enhanced its resolution through a Schur complement. Numerical experiments give satisfying results.

Outlooks:

- how to choose N_{ex} ? $\rightsquigarrow$ estimate the conditioning of the Schur complement to reach a given enhancement of the convergence, but still requires a first SCF calculation.
- how to adapt on the fly (*i.e.* during SCF) the number of extra bands ?
- implemented by default in DFTK, which allows for efficient Automatic Differentiation implementations.

Thanks for your attention !

Joint work with

Éric Cancès

Michael F. Herbst

Antoine Levitt

Benjamin Stamm