Using the density-functional toolkit (DFTK)

to investigate floating-point error and SCF convergence of mixed systems

Michael F. Herbst, Antoine Levitt, Eric Cancès

Matherials team, CERMICS, École des Ponts ParisTech

09th January 2020

→ `https://michael-herbst.com/talks/2020.01.09_dftk_fp_scf_lille.pdf`

Contents

Screening applications of DFT

Materials and semiconductors

CC-by-3.0 https://en.wikipedia.org/wiki/File:Silicon.jpg

Chemical and pharmaceutical industry

CC-by-1.0 https://en.wikipedia.org/wiki/File:Lab_bench.jpg

- Interest in high-throughput screening
  - Drug design, catalysis, material science, . . .
- As fast as possible, as accurate as necessary
- Flexible and reliable workflows . . .

Screening applications of DFT

Materials and semiconductors

CC-by-3.0 https://en.wikipedia.org/wiki/File:Silicon.jpg

Chemical and pharmaceutical industry

CC-by-1.0 https://en.wikipedia.org/wiki/File:Lab_bench.jpg

- Interest in high-throughput screening
  - Drug design, catalysis, material science, ...
- As fast as possible, as accurate as necessary
- Flexible and reliable workflows ...

Existing codes

- `https://www.vasp.at`
- `https://www.abinit.org`
- `http://www.castep.org`
- `https://wiki.fysik.dtu.dk/gpaw`
- `https://www.quantum-espresso.org`
- `https://crd-legacy.lbl.gov/~chao/KSSOLV`
- ...

⇒ Why another? Why julia?

Questions related to high-throughput

- As fast as possible, as accurate as needed
  - Effect of floating-point type?
  - Dynamically raise / lower precision?
  - Control floating-point error?
  - High-performance computing: GPU, distributed ...
- Flexible and reliable workflows
  - Reliable, black-box SCF algorithms
  - Difficult for mixed systems!
  - Small and understandable code base
  - High-level API and dynamic language

$\Rightarrow$ Interdisciplinary setting

Demands for interdisciplinary software

- **Mathematicians:** Toy models and unphysical edge cases
- **Scientist:** Wants to focus on science, not numerics
- **High-performance person:** Exploit hardware specialities
- **Practitioner:** Reliable, black-box, high-level interface

- julia: Our language of choice
  - *Walks like Python, talks like Lisp, runs like FORTRAN*
  - High-level, but hackable language (on all levels!)
  - Rich ecosystem, support for AD, GPUs, special hardware

  $\Rightarrow$ Write code once, re-use for many back ends / machines ...

- `https://michael-herbst.com/learn-julia`

DFTK — https://dftk.org

- 8 months of development, < 3000 lines
- Accessible to interdisciplinary community:
  - Custom Hamiltonians, potentials
  - Construct new models on a high-level
  - Support for arbitrary floating point types
  - Target: Modern HPC environments
- (Most) standard ground-state features:
  - LDA, GGA functionals from `libxc`, GTH pseudopotentials
  - SCF (DIIS, NLsolve, damping) or direct minimisation
  - Smearing, Kerker preconditioner for metals
  - Shared-memory parallelism
  - Tested against ABINIT (for three cases)

DFTK — https://dftk.org

- 8 months of development, < 3000 lines
- Accessible to interdisciplinary community:
  - Custom Hamiltonians, potentials
  - Construct new models on a high-level
  - Support for arbitrary floating point types
  - Target: Modern HPC environments
- (Most) standard ground-state features:
  - LDA, GGA functionals from `libxc`, GTH pseudopotentials
  - SCF (DIIS, NLsolve, damping) or direct minimisation
  - Smearing, Kerker preconditioner for metals
  - Shared-memory parallelism
  - Tested against ABINIT (for three cases)

Contents

Density-functional theory (Euler-Lagrange equations)

$$\begin{cases} \underbrace{\left(-\frac{1}{2}\Delta + V_{\text{Nuc}} + V_{\text{H}}\left[\rho_{\Phi_0}\right] + V_{\text{xc}}\left[\rho_{\Phi_0}\right]\right)}_{=\hat{\mathcal{F}}_{\Phi_0}} \psi_i = \varepsilon_i \psi_i, \\ \langle \psi_i | \psi_j \rangle = \delta_{ij}, \psi_i \in H^1(\mathbb{R}^3, \mathbb{C}) \qquad \forall i, j \quad 1 \leq i, j \leq N \\ -\Delta V_{\text{H}}\left[\rho_\Phi\right] = 4\pi\rho_\Phi \\ V_{\text{xc}}[\rho_\Phi]\,\mathrm{d}\rho_\Phi = \mathrm{d}\left(\int_{\mathbb{R}^3} \epsilon_{\text{xc}}[\rho_\Phi](\underline{\boldsymbol{r}})\,\mathrm{d}\underline{\boldsymbol{r}}\right) \qquad \text{(Fréchet sense)} \end{cases}$$

with Fermi-Dirac distribution $f\varepsilon_F$, density

$$\rho_\Phi(\underline{\boldsymbol{r}}) = f_{\varepsilon_F}\left(\hat{\mathcal{F}}_\Phi\right)(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}),$$

and Fermi level ε_F s.t. $\int_{\underline{\boldsymbol{r}}} \rho_\Phi(\underline{\boldsymbol{r}})\,\mathrm{d}\underline{\boldsymbol{r}} = N$

Density-functional theory (Euler-Lagrange equations)

$$\begin{cases} \underbrace{\left(-\frac{1}{2}\Delta + V_{\text{Nuc}} + V_{\text{H}}\left[\rho_{\Phi_0}\right] + V_{\text{xc}}\left[\rho_{\Phi_0}\right]\right)}_{=\hat{\mathcal{F}}_{\Phi_0}} \psi_i = \varepsilon_i \psi_i, \\ \langle \psi_i | \psi_j \rangle = \delta_{ij}, \psi_i \in H^1(\mathbb{R}^3, \mathbb{C}) \qquad \forall i,j \quad 1 \leq i,j \leq N \\ -\Delta V_{\text{H}}\left[\rho_\Phi\right] = 4\pi\rho_\Phi \\ V_{\text{xc}}[\rho_\Phi]\,\mathrm{d}\rho_\Phi = \mathrm{d}\left(\int_{\mathbb{R}^3} \epsilon_{\text{xc}}[\rho_\Phi](\underline{\boldsymbol{r}})\,\mathrm{d}\underline{\boldsymbol{r}}\right) \qquad \text{(Fréchet sense)} \end{cases}$$

with Fermi-Dirac distribution $f\varepsilon_F$, density

$$\rho_\Phi(\underline{\boldsymbol{r}}) = f_{\varepsilon_F}\left(\hat{\mathcal{F}}_\Phi\right)(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}),$$

and Fermi level ε_F s.t. $\int_{\underline{\boldsymbol{r}}} \rho_\Phi(\underline{\boldsymbol{r}})\,\mathrm{d}\underline{\boldsymbol{r}} = N$

SCF and errors

- Self-consistency problem:

$$\text{Guess } \Phi \;\Rightarrow\; \text{Compute } \hat{\mathcal{F}}_\Phi \;\Rightarrow\; \text{Obtain new } \Phi$$

- Discretisation: Plane-wave basis at each $\underline{\boldsymbol{k}}$-Point:

$$\left\{ \frac{e^{\imath \underline{\boldsymbol{G}} \cdot \underline{\boldsymbol{r}}}}{\sqrt{\|\underline{\boldsymbol{G}}\|}} \;\middle|\; \underline{\boldsymbol{G}} \in \mathcal{R}^*, \|\underline{\boldsymbol{G}} + \underline{\boldsymbol{k}}\|^2 \leq 2E_{\text{cut}} \right\}$$

- Sources of errors:
  - Method / DFT functional (approximate model)
  - Discretisation insufficient
  - SCF procedure not completely converged
  - Floating-point arithmetic inexact

Determination of errors

- Exact quantity q, current iterate $q^{(i)}$
- Error: $e^{(i)} = \left\| q - q^{(i)} \right\|_V$
  - Examples: $\left\| \rho - \rho^{(i)} \right\|_V$, $|E - E^{(i)}|$
- SCF residual:

$$r^{(i)} = \hat{\mathcal{F}}_{\Phi^{(i)}} \Phi^{(i)} - \Lambda^{(i)} \Phi^{(i)}$$

- Numerical analysis: Find constant C and norms W, s.t.

$$e^{(i)} \leq C \left\| r^{(i)} \right\|_W \qquad \text{(\textit{a posteriori} error)}$$

- Choice of norms: $\|\cdot\|_\infty$ $\|\cdot\|_2$ $\|\cdot\|_{H^1}$ $\|\cdot\|_{H^{-1}}$
- Norm V depends on the property of interest

Determination of errors

- Exact quantity q, current iterate $q^{(i)}$
- Error: $e^{(i)} = \left\| q - q^{(i)} \right\|_V$
  - Examples: $\left\| \rho - \rho^{(i)} \right\|_V$, $|E - E^{(i)}|$
- SCF residual:

$$r^{(i)} = \hat{\mathcal{F}}_{\Phi^{(i)}} \Phi^{(i)} - \Lambda^{(i)} \Phi^{(i)}$$

- Numerical analysis: Find constant C and norms W, s.t.

$$e^{(i)} \leq C \left\| r^{(i)} \right\|_W \qquad (\textit{a posteriori} \text{ error})$$

- Choice of norms: $\|\cdot\|_\infty$ $\|\cdot\|_2$ $\|\cdot\|_{H^1}$ $\|\cdot\|_{H^{-1}}$
- Norm V depends on the property of interest

Finding the right norms is not always easy ...

- Consider a linear problem

$$\mathbf{A}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$$

- Oettli-Prager relation:

$$\left\|\underline{\boldsymbol{x}}^{(i)} - \underline{\boldsymbol{x}}\right\|_\infty \leq \frac{\left\|\mathbf{A}\underline{\boldsymbol{x}}^{(i)} - \underline{\boldsymbol{b}}\right\|_\infty}{\|\mathbf{A}\|_\infty \left\|\underline{\boldsymbol{x}}^{(i)}\right\|_1 + \|\underline{\boldsymbol{b}}\|_\infty}$$

Characterising floating-point error

- Ignore method error, discretisation error
- For a converged Φ:

$$0 = r = \hat{\mathcal{F}}_\Phi \Phi - \Lambda \Phi$$

- In practice: $r \neq 0$

$\Rightarrow$ Inexact floating-point arithmetic

- From converged Φ: Measure floating-point error
- Verifies whether Φ is meaningful!

Interval arithmetic

- Represent $x \in \mathbb{R}$ by the interval

$$[a, b] \quad a, b \in \mathrm{DP}, \quad a \leq x \leq b$$

- Computations done *simultaneously* on a and b, e.g.

$$\exp([a, b]) = [\exp(a), \exp(b)]$$

- More complicated for non-monotone functions ...
- Lower / Upper limited rounded down / up
- Result interval contains exact answer

$\Rightarrow$ Speed: Fourier-transforms ≈ 100 times slower

DEMO

DEMO

Numerical error in SCF residual

Contents

SCFs as a fixed-point problem

- Self-consistency problem: Fixed-point problem in ρ

$$\rho = F(\mathcal{V}(\rho))$$

- Corresponding Jacobian is

$$J = 1 - \chi_0(v_C + K_{\text{xc}})$$

  - v_C: Periodised Coulomb kernel
  - K_{xc}: Exchange-correlation kernel
  - χ_0: Independent-particle susceptibility

- Mixing schemes: Quasi-Newton updates

$$\rho_{\text{next}} = \rho_{\text{in}} + C\,(\rho_{\text{out}} - \rho_{\text{in}})$$

where $C \approx J^{-1}$ and $\rho_{\text{out}} = F(\mathcal{V}(\rho_{\text{in}}))$.

Typical mixing schemes

$$\rho_{\text{next}} = \rho_{\text{in}} + C\left(\rho_{\text{out}} - \rho_{\text{in}}\right)$$

- Insulators: Simple mixing $C = 0 < \alpha \ll 1$
- Metals: Kerker mixing $(\alpha, \gamma > 0)$

$$C = \frac{-\alpha\Delta}{\gamma - \Delta}, \qquad \hat{C} = \frac{\alpha q^2}{q^2 + 4\pi\gamma}$$

- Prevent charge sloshing:
  - Low-frequency, large-wavelength components

$\Rightarrow$ For hybrid systems?

Hybrid mixing

- Idea: Approximate the χ_0 in

$$J^{-1} = \left(1 - \chi_0(v_C + K_{\text{xc}})\right)^{-1}$$

- May obtain J^{-1} iteratively (solving a linear system)

$\Rightarrow$ Key step is to apply J, thus apply χ_0:

$$(\chi_0\phi)(\underline{\boldsymbol{r}}) = \int_\Gamma \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}')\phi(\underline{\boldsymbol{r}}')\,\mathrm{d}\underline{\boldsymbol{r}}' \simeq \phi(\underline{\boldsymbol{r}}) \int_\Gamma \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}} + \underline{\boldsymbol{r}}')\,\mathrm{d}\underline{\boldsymbol{r}}'$$

Hybrid mixing

- Idea: Approximate the χ_0 in

$$J^{-1} = \left(1 - \chi_0(v_C + K_{\text{xc}})\right)^{-1}$$

- May obtain J^{-1} iteratively (solving a linear system)

$\Rightarrow$ Key step is to apply J, thus apply χ_0:

$$(\chi_0\phi)(\underline{\boldsymbol{r}}) = \int_\Gamma \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}}')\phi(\underline{\boldsymbol{r}}')\,\mathrm{d}\underline{\boldsymbol{r}}' \simeq \phi(\underline{\boldsymbol{r}}) \int_\Gamma \chi_0(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}} + \underline{\boldsymbol{r}}')\,\mathrm{d}\underline{\boldsymbol{r}}'$$

DEMO

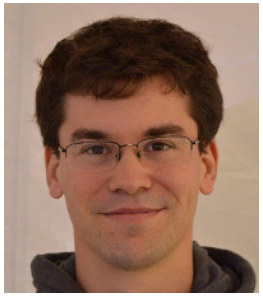
DEMO

SCF convergence in hybrid systems

Summary and outlook

- `DFTK`: Interdisciplinary software development
  - Concise, feature-rich julia toolkit
  - Toy problems *and* full-scale applications
- Floating-point error in DFT
  - Determine and understand floating-point error in DFT
  - Problem-specific refinement of arithmetic *and* discretisation
- SCF hybrid mixing
  - Preconditioning based on local density of states
  - Tests for challenging realistic systems (spin?)

Acknowledgements



Antoine Levitt

Eric Cancès

Inria

École des Ponts ParisTech

SORBONNE UNIVERSITÉ

European Research Council
Established by the European Commission

Questions?

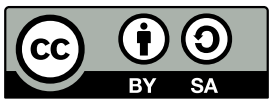
DFTK: https://dftk.org

julia: https://michael-herbst.com/learn-julia

 mfherbst

 https://michael-herbst.com/blog

 michael.herbst@inria.fr



Definitions

- Potential-to-density map F

$$F(V) = f_{\varepsilon_F}\left(-\frac{1}{2}\Delta + V\right)(\underline{\boldsymbol{r}}, \underline{\boldsymbol{r}})$$

- Density-to-potential map $\mathcal{V}(\rho)$
- Fixed-point problem

$$\rho = F(\mathcal{V}(\rho))$$

with residual function

$$r(\rho) = \rho - F(\mathcal{V}(\rho))$$

and Jacobian (derivative of residual)

$$J = 1 - \chi_0(v_C + K_{\text{xc}})$$

- K_{xc} usually smaller eigvals than v_C, thus ignored.

Spectral properties of Jacobian

- Newton scheme:

$$\rho^{(i+1)} = \rho^{(i)} + J^{-1}\left(F(\mathcal{V}(\rho^{(i)})) - \rho^{(i)}\right)$$

- Typically J^{-1} approximated (Quasi-Newton):

$$\rho^{(i+1)} = \rho^{(i)} + C^{(i)}\left(F(\mathcal{V}(\rho^{(i)})) - \rho^{(i)}\right)$$

- Near the fixed point:

$$\delta\rho^{(i+1)} \approx \delta\rho^{(i)} - C^{(i)}J^*\delta\rho^{(i)} = (I - C^{(i)}J^*)\delta\rho^{(i)},$$

  where J^* is Jacobian at fixed-point.

$\Rightarrow$ $\sigma\left((I - C^{(i)}J^*)\right) < 1$ desirable

$\Rightarrow$ $0 < \sigma(C^{(i)}J^*) < 2$

Metals versus insulators

- Ignore XC term: $J \simeq 1 - \chi_0 v_C$
- For charge-sloshing effects, we care about low q regime:
  - Low frequency, large wavelength density response
  - Connects remote density areas (not physical)
- v_C goes as $4\pi/q^2$
- For metals $\chi_0 \ -\gamma$ for $\gamma > 0$ for small q, therefore J^{-1} is approximately Kerker:

$$\frac{1}{1 + 4\pi\gamma/q^2} = \frac{q^2}{q^2 + 4\pi\gamma}$$

- For insulators $\chi_0 \ -\xi q^2$ for $\xi > 0$: $\sigma(J)$ bounded

  $\Rightarrow$ $C^{(i)} = \alpha$ for small enough α is sufficient.