

Density functional theory in solid state physics

Lecture 8

- Planewave basis useful and very accurate for periodic solids
- localized basis useful for low-D systems, potentially faster than PW, but more complex
- Kohn-Sham Equation in reciprocal space (solve iteratively):

$$\sum_{\vec{G}'} \left[\frac{1}{2} |\vec{k} + \vec{G}|^2 \delta_{\vec{G}\vec{G}'} + v(\vec{G}, \vec{G}') + v_H(\vec{G} - \vec{G}') + v_{XC}(\vec{G} - \vec{G}') \right] c_{ik}(\vec{G}') = \epsilon_{ik} S c_{ik}(\vec{G})$$

Describe this with pseudo- or PAW potentials.

Non-local PPs: $v_k^{NL-PS}(\vec{G}, \vec{G}')$

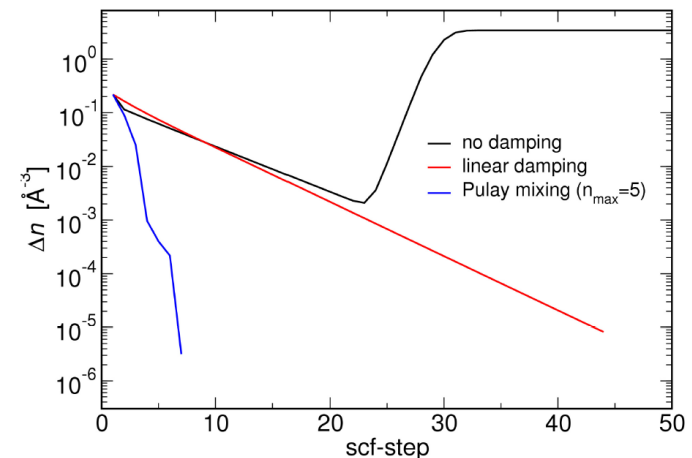
Want to only consider valence electrons

Overlap matrix,
For NC-PP: $S = I$
For US: $S = I + \sum_{mn} Q_{mn}(\vec{r}) |\beta_m\rangle \langle \beta_n|$

- Solution using SCF algorithm can suffer from charge sloshing
- Need to stabilize SCF: Mix new and old electron densities

$$n_{in}^i(r) = \alpha n_{out}^{i-1}(r) + (1 - \alpha) n_{in}^{i-1}(r)$$

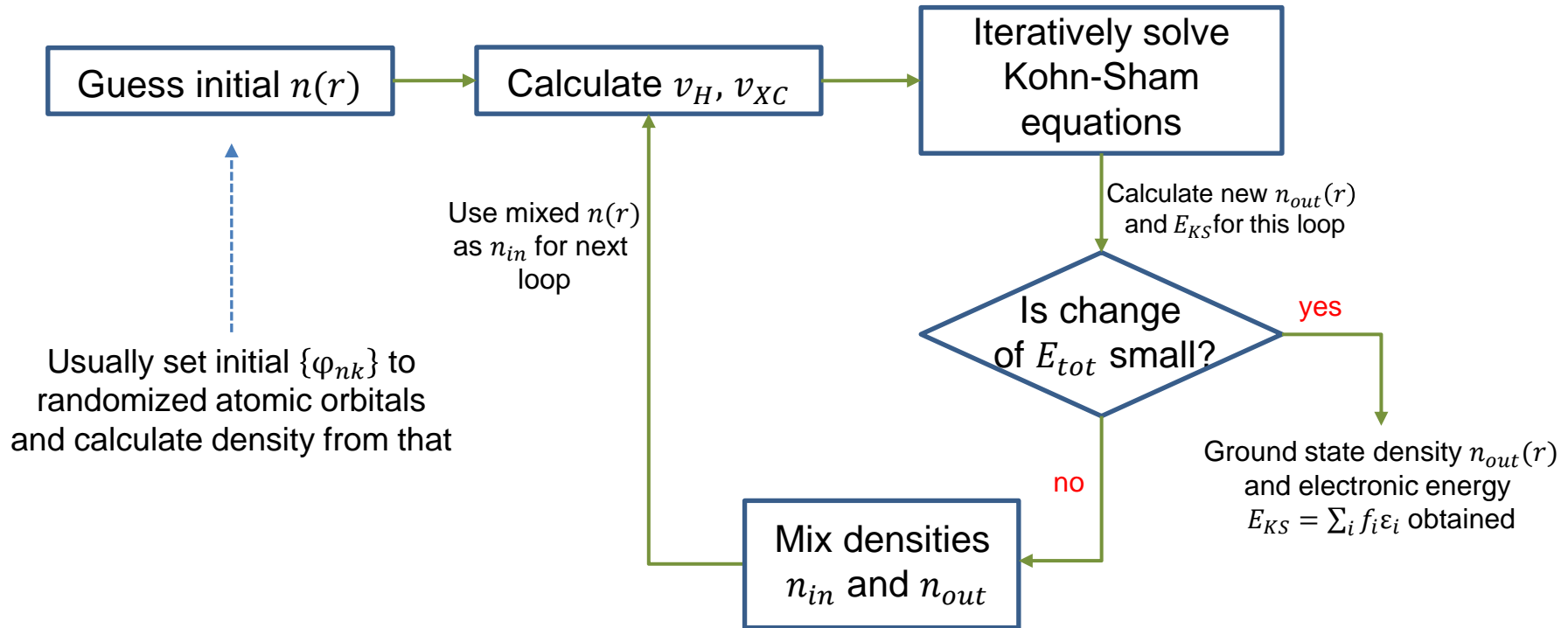
mixing parameter



Taken from „Making electronic structure theory work“

The SCF method with density mixing

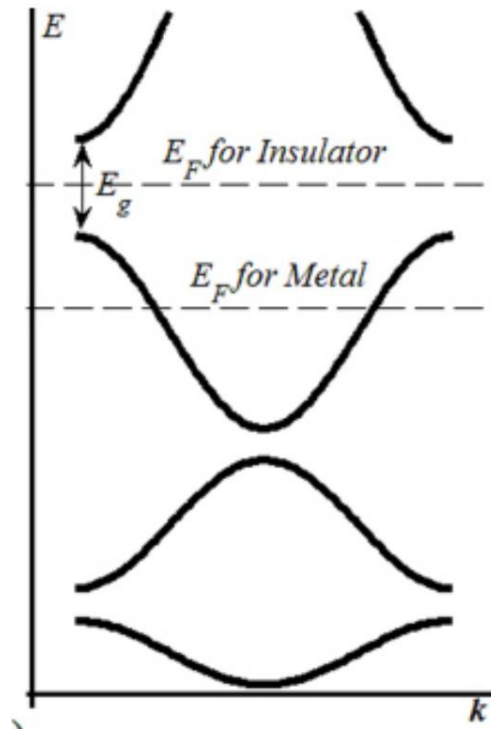
- More stable SCF algorithm:



- Alternatively also can mix Kohn-Sham potential $v_{KS}(r)$, but less stable than density mixing

Charge sloshing for metals

- Metals: partial occupation of bands.
- Slow convergence with k-point sampling due to integration of discontinuous occupation
- Small density variations can cause strong changes in band occupations



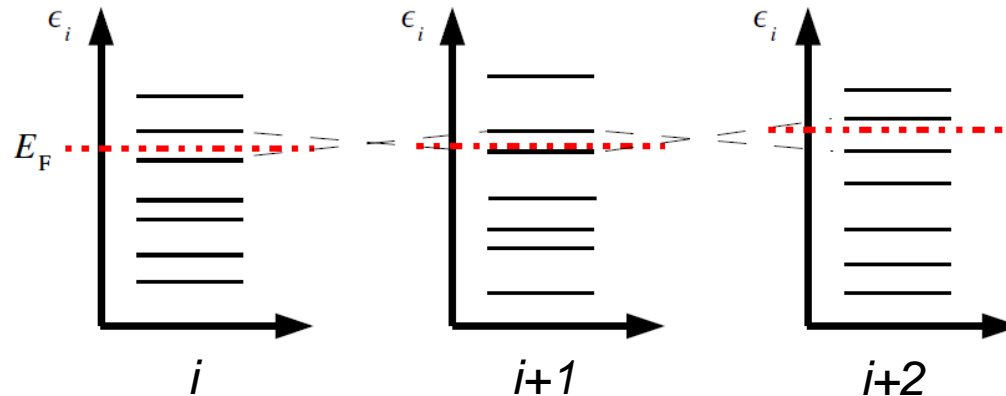
$$n(r) = \frac{1}{V_{BZ}} \sum_n \int_{BZ} f(\epsilon_{nk} - E_F) |\varphi_{nk}|^2 dk$$

$$E_{KS} = \frac{1}{V_{BZ}} \sum_n \int_{BZ} f(\epsilon_{nk} - E_F) \epsilon_{nk} dk$$

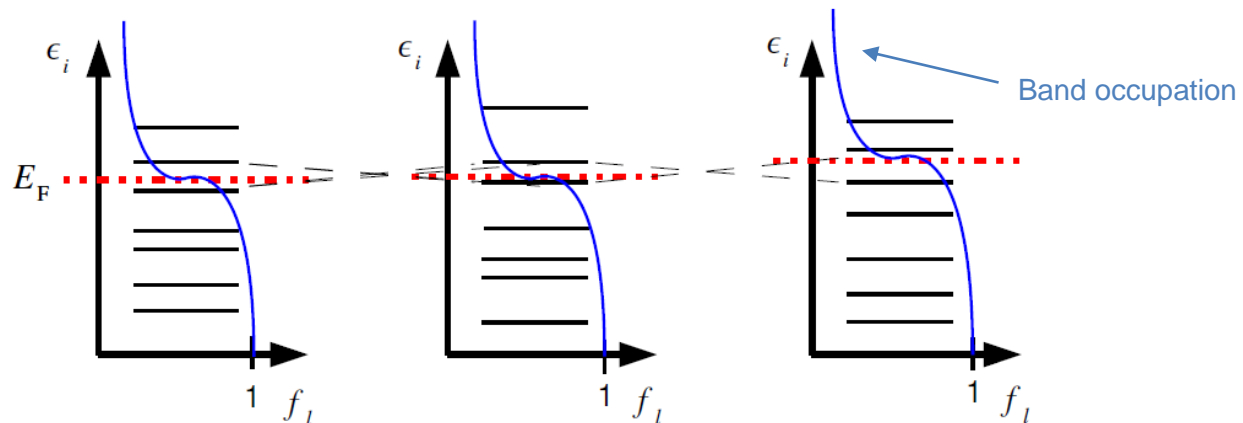
$$f(x) = \theta(x) = \begin{cases} 1, & \text{if } x \leq 0 \\ 0, & \text{if } x > 0 \end{cases}$$

Level crossings in metals

- Additional problem: discontinuous occupation leads to „jumps“ in energy, level crossings



- Remedy: „Smear out“ bands occupations around the Fermi energy

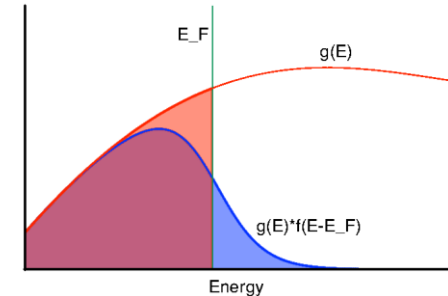


Taken from „Making electronic structure theory work“

Smearing of band occupations

- Electron density: $n(r) = \sum_{n,k} f(\epsilon_{nk}) |\varphi_{nk}(r)|^2$
- Fermi-Dirac smearing for „electronic temperature“ T_e :

$$f_{FD}(\epsilon_{nk}) = \left[\exp\left(\frac{\epsilon_{nk} - E_F}{\sigma}\right) + 1 \right]^{-1} \quad \sigma = k_B T_e$$



Introduction of error, as Fermi-Dirac is a bad approximation to a step function, slow convergence with respect to σ

- Gaussian smearing: Approximate step function (occupation at $T = 0$) by smooth gaussian

$$f_{gauss}(\epsilon_{nk}) = \frac{1}{2} \left(1 - \operatorname{erf} \left(\frac{\epsilon_{nk} - E_F}{\sigma} \right) \right)$$

Energy width of smearing

Faster convergence with k-point grid than Fermi-Dirac, but not guaranteed that SCF converges to correct minimum

Still: standard smearing method in most DFT solid-state physics codes

- Methfessel-Paxton: Approximate step function by expansion in Hermite polynomials

$$f_{MP,0}(x) = \frac{1}{2} (1 - \operatorname{erf}(x))$$
$$f_{MP,N}(x) = f_{MP,0}(x) + \sum_{n=1}^N A_n H_{2n-1}(x) e^{-x^2}$$
$$x = \frac{\varepsilon_{nk} - E_F}{\sigma}$$

Gaussian smearing is equal to 0-order MP smearing, order-1 and order-2 are most widely used

Very fast convergence with σ and k-point sampling, but might get unphysical *negative* band occupations

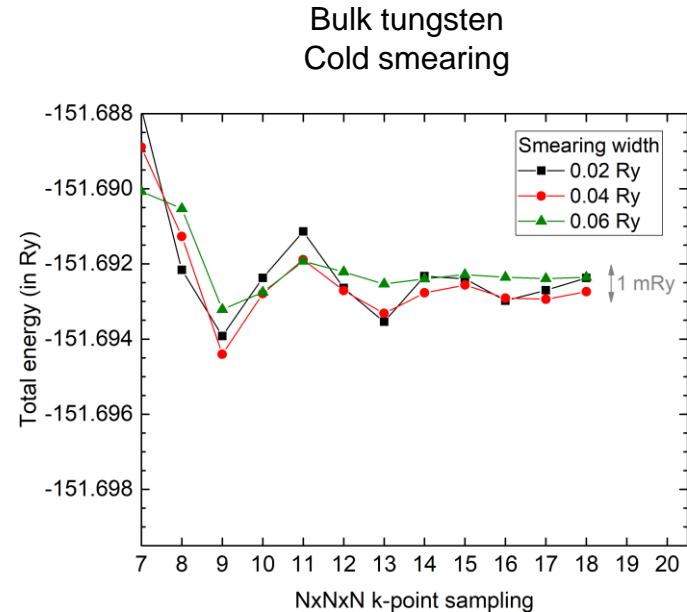
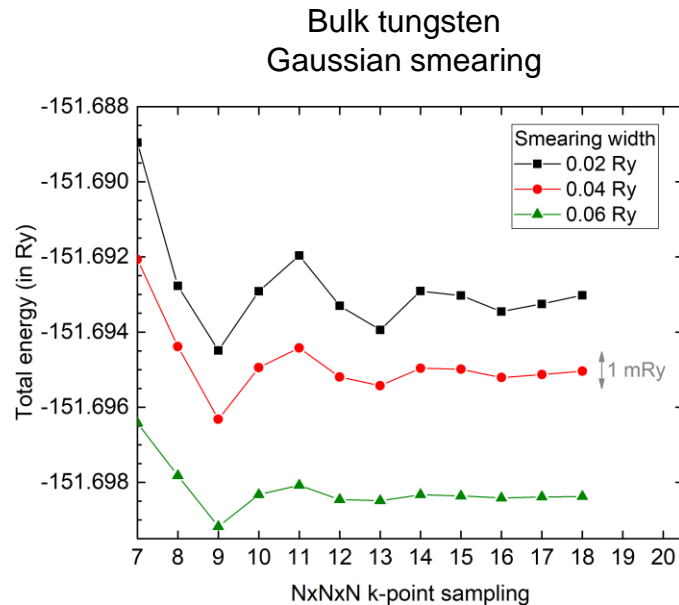
Phys. Rev. B 40, 3616 (1989)

- Marzari-Vanderbilt („cold smearing“): Generalization of order-2 MP smearing to always yield positive occupations


(Phys. Rev. Lett. 82, 3296 (1999))

Total energy dependence on smearing

- Obtained energy depends on smearing width/electron temperature

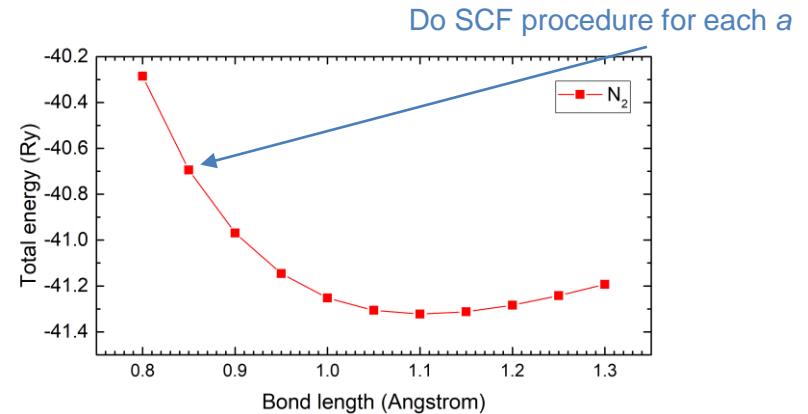


- Variational minimization from SCF not total energy anymore, but free energy $F = E - \sigma S$
- Need to back-extrapolate to $E[\sigma \rightarrow 0]$ in order to obtain desired E_{tot}

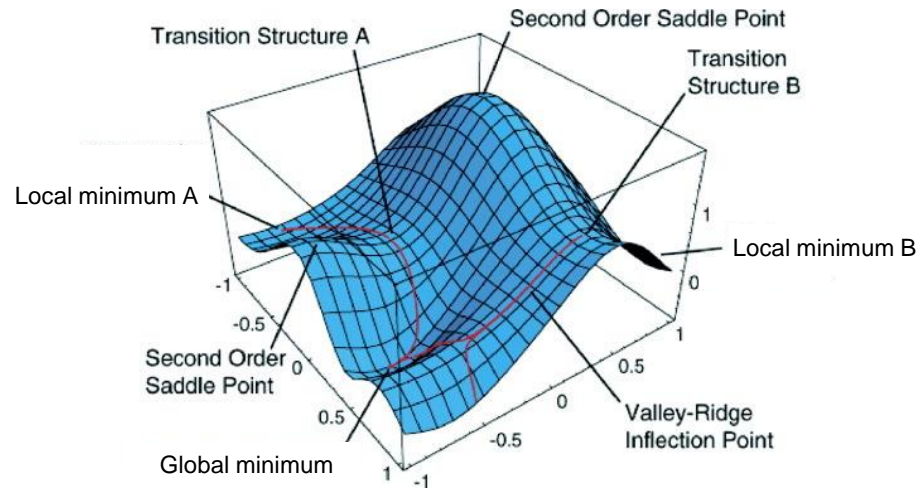
- So far: solution of Kohn-Sham Equations for a fixed set of atomic positions
- Direct access to a variety of properties:
 - Total energy $E_{tot} = E_{KS} + E_{n-n}$
(comparison with other structures, binding energy of adsorbates etc)

 - Electron density $n(r)$
 - Potentials v, v_H, v_{XC} and $v_{KS} = v + v_H + v_{XC}$
 - Quasi-single electron eigenvalues ε_{nk} and band occupations f_{nk}
 - Fermi energy E_F from band occupations (and with that also work function)
- Properties will somewhat depend on used atomic positions
- To be fully consistent, also need to take nucleic degree of freedom into account
- Also: How do we get the atomic positions and lattice constants for materials without tabulated structural data?

Manual minimization of atomic positions

- Geometry optimization by varying atomic positions and minimizing total energy E_{tot}
- Simple example: diatomic molecule



- For more complicated systems: Find global minimum of potential energy surface:



→ complex problem

Hellmann-Feynman forces

- Classically, driving forces acting on nucleus l : $\vec{F}_l = -\frac{\partial E_{tot}}{\partial \vec{R}_l}$
- Energy minimum corresponds to vanishing interatomic forces: $\vec{F}_l = 0|_{\{\vec{R}_l\}=\{\vec{R}_l^0\}}$ for all l
- Hellmann-Feynman theorem for perturbation $\partial \vec{R}_l$:

$$\frac{\partial E_{tot}}{\partial \vec{R}_l} = \frac{\partial}{\partial \vec{R}_l} \langle \psi | \hat{H} | \psi \rangle = \left\langle \psi \left| \frac{\partial \hat{H}}{\partial \vec{R}_l} \right| \psi \right\rangle + \underbrace{E_{tot} \frac{\partial}{\partial \vec{R}_l} \langle \psi | \psi \rangle}_{\text{"Pulay forces"}}$$

Many-body wavefunction

These vanish, if basis set is independent from nucleic positions (such as planewaves)

- In DFT:

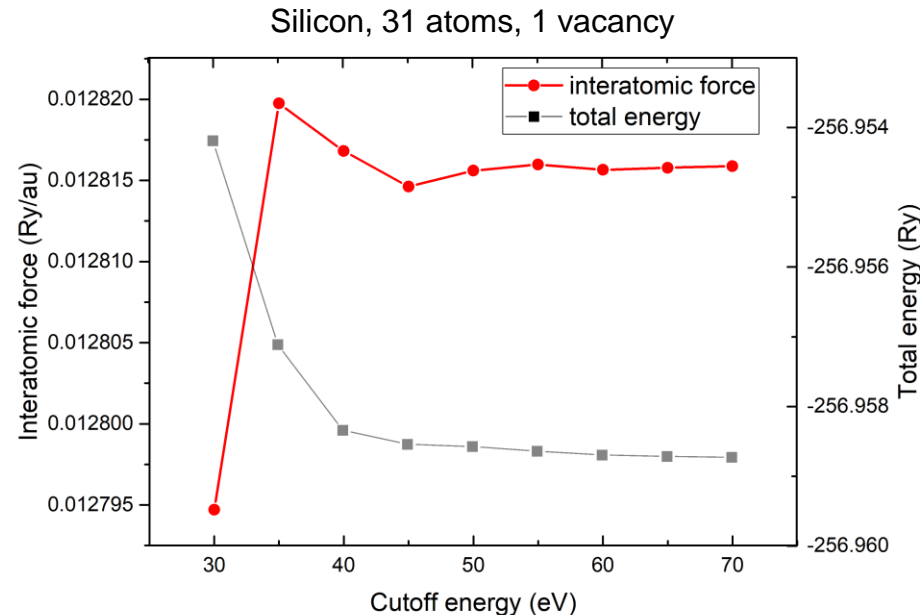
$$E_{tot}[\{R\}] = E_{KS}[\{R\}] + E_{n-n}[\{R\}]$$

$$\rightarrow \frac{\partial E_{tot}}{\partial \vec{R}_l} = \frac{\partial E_{n-n}}{\partial \vec{R}_l} - \int n(r) \frac{\partial v}{\partial \vec{R}_l} dr - \int \frac{\partial E_{KS}}{\partial n} \frac{\partial n}{\partial \vec{R}_l} dr$$

This term vanishes in the energy minimum

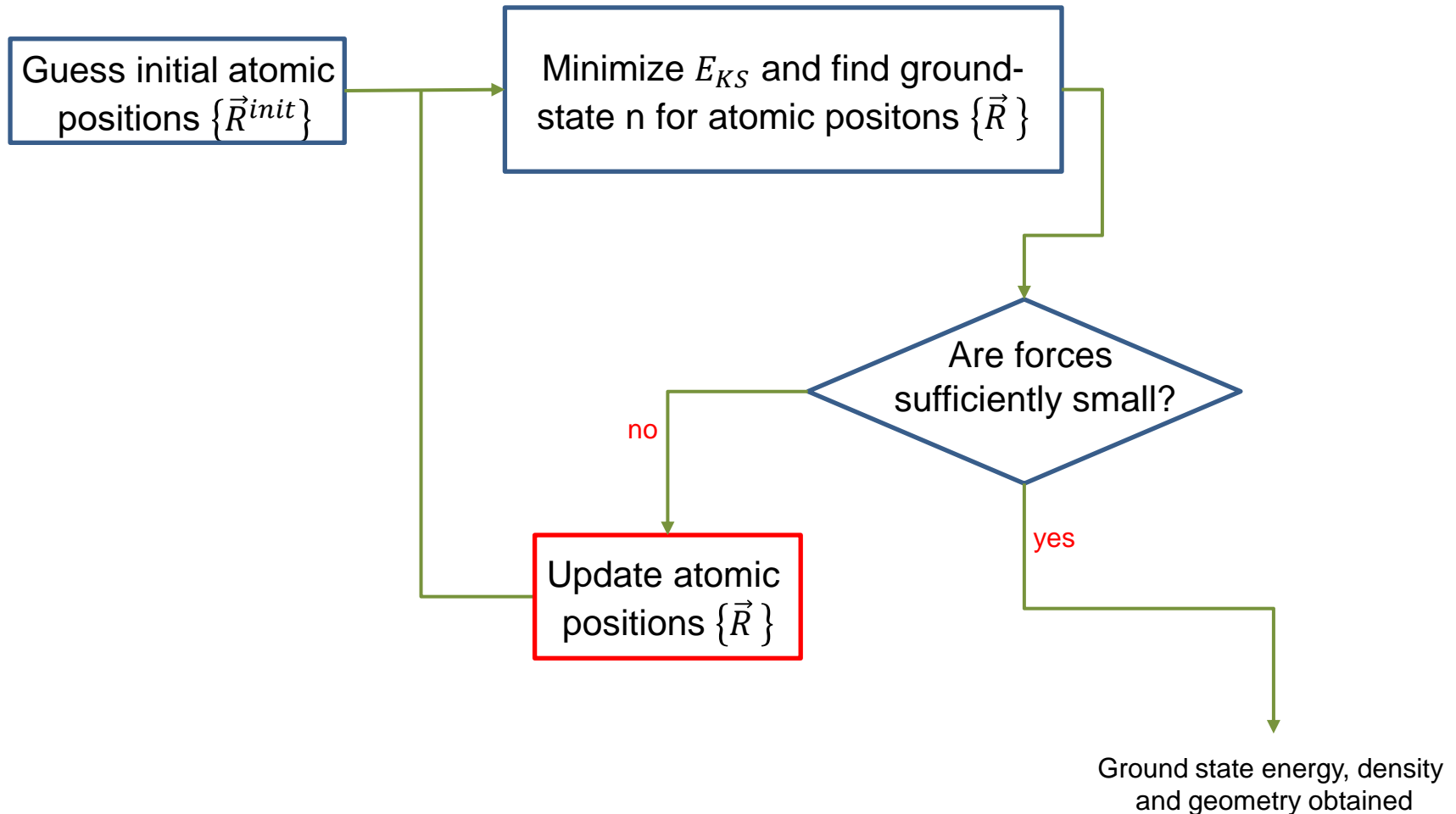
Convergence of interatomic forces

- Total energy from *variational* SCF procedure correct to second order errors, but forces only correct to first order errors
- More complicated, if we use density mixing to solve KS equations
- Careful: Convergence of forces with basis set size (cutoff energy) is (usually) slower than that of E_{tot}
- Should also check influence of energy tolerance for SCF convergence (Exercise sheet 2)



Iterative geometry optimization

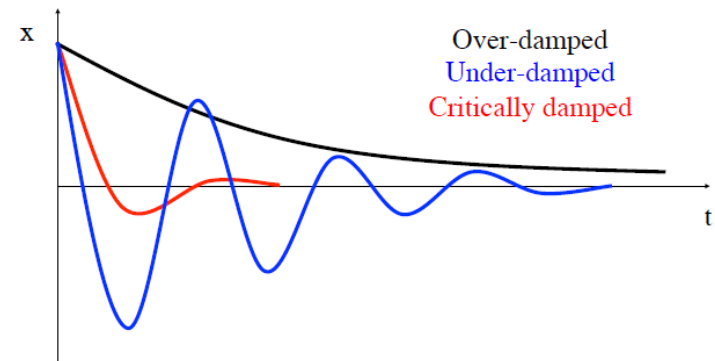
- Treat geometry optimization as an iterative process:



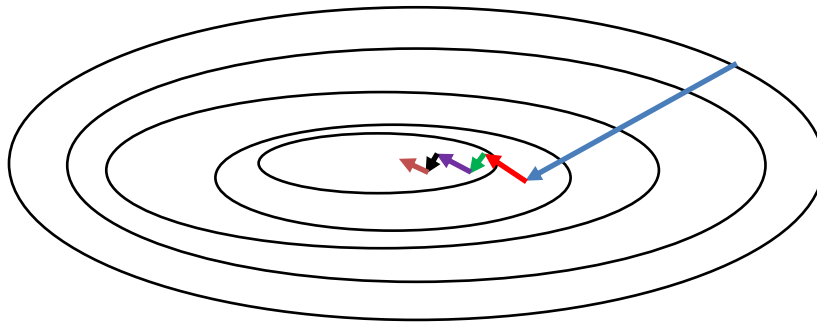
- How do we update atomic positions?
- Physically motivated approach: damped molecular dynamics (damped MD):
 - Treat atoms as damped harmonic oscillators

$$M_l \ddot{\vec{R}}_l = \vec{F}_l - \gamma \vec{v}_l$$

- Move atoms using obtained velocities and forces, need to choose suitable time step Δt
- Use an algorithm to guess optimal damping γ
- Fast convergence, while far away from minimum, but slower convergence when close to minimum (small forces)



- Alternatively: treat update of atomic positions as a iterative minimization problem
- Steepest descent (SD)



1. Line search: Move atoms along energy gradient (i.e. forces) until local energy starts to rise again
2. Calculate new gradient and repeat 1. until forces are small enough

- convergence of steepest descent towards minimum is unpractically slow
- Conjugate gradients (CG)
 - Similar to SD, but choose new search directions to be orthogonal to all previous search directions
 - Significantly faster convergence than SD

- The Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm
 - Quasi-Newton method
 - Idea: energy surface near the minimum is nearly quadratic, i.e. determined by the Hessian matrix \vec{A} of the system

$$\delta E_{tot} = \frac{1}{2} (\vec{R} - \vec{R}^{min})^T \cdot \vec{A} \cdot (\vec{R} - \vec{R}^{min}) \quad \vec{A} = \begin{pmatrix} \frac{\partial^2 E_{tot}}{\partial^2 \vec{R}_1} & \dots & \frac{\partial^2 E_{tot}}{\partial \vec{R}_1 \partial \vec{R}_N} \\ \vdots & \ddots & \vdots \\ \dots & \dots & \frac{\partial^2 E_{tot}}{\partial^2 \vec{R}_N} \end{pmatrix}$$

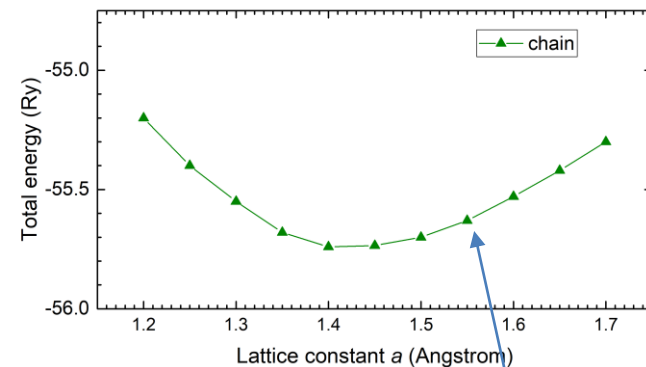
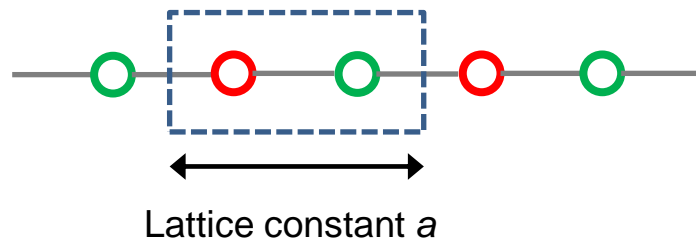
- If Hessian known, we could find the minimum in one step
- Algorithm:
 1. Guess approximate Hessian
 2. Do a line search along $\Delta \vec{X}_i = (\vec{A}_i)^{-1} \vec{F}_i$ and update atomic positions: $\vec{X}_{i+1} = \vec{X}_i + \lambda \Delta \vec{X}_i$
 3. Improve approximation of Hessian and repeat step 2 until convergence achieved
- BFGS is far more efficient than CG or SD, standard method in most codes
- Careful: iterative methods can be caught in local minima instead of finding global minimum

Matrix of interatomic
forces for step i

Matrix of atomic
positions for step i

Manual update of lattice vectors

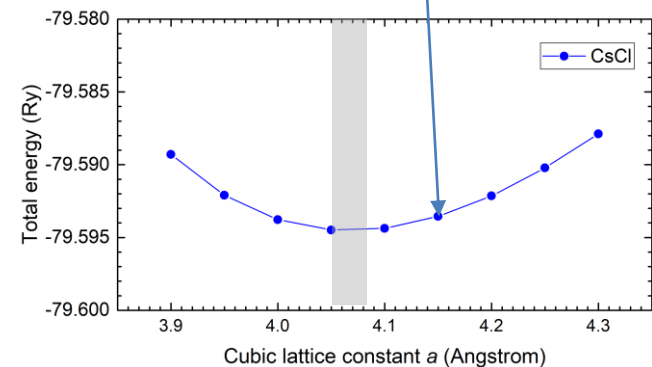
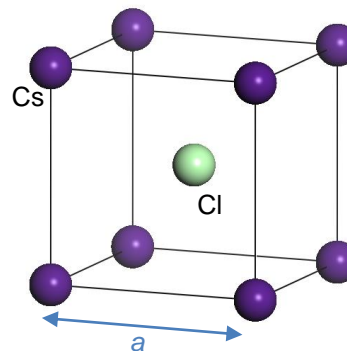
- How about lattice constants/vectors?
- In principle could again update manually
- Simple example: Atomic chain



Optimize atomic positions for each a

- Similarly for more complex 3D systems

Example: bcc CsCl



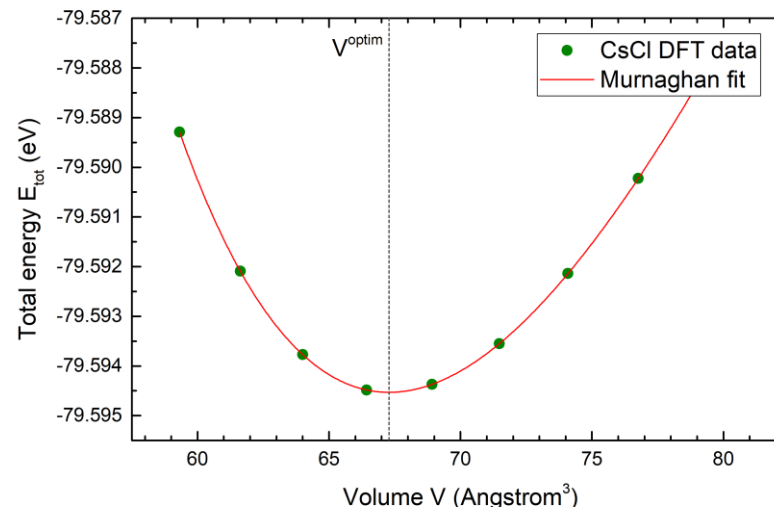
Birch-Murnaghan Equation-Of-State

- A bit better: Fit Birch-Murnaghan Equation to DFT data

$$E_{tot}(V) = E_0 + \frac{9V_0B_0}{16} \left\{ B'_0 \left[\left(\frac{V_0}{V} \right)^{2/3} - 1 \right]^3 + \left[6 - 4 \left(\frac{V_0}{V} \right)^{2/3} \right] \left[\left(\frac{V_0}{V} \right)^{2/3} - 1 \right]^2 \right\}$$

Groundstate total energy

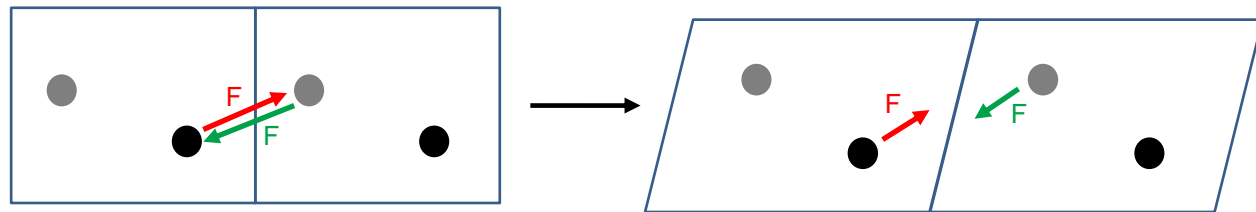
V_0 : Unit cell volume at energy minimum, B_0 : Bulk modulus, B'_0 : Derivative of bulk modulus with pressure



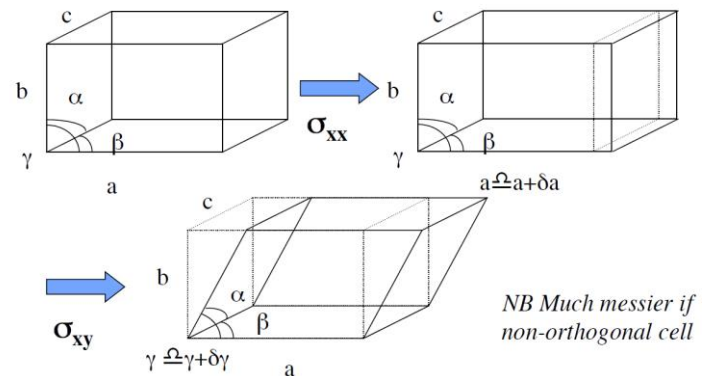
- Advantage of BM method: Can directly obtain lattice constants and bulk moduli B_0
- However: the two methods are (obviously) difficult to use for Bravais lattices with low symmetry (i.e. with different lattice constants and angles)

Variable cell relaxations

- Common approach nowadays: „Variable cell“ optimizations
- Situation: Interatomic forces between neighbouring cells cause stresses $\sigma = \frac{F}{A}$ on the cell faces



- Stress induces volume changes and shape deformations

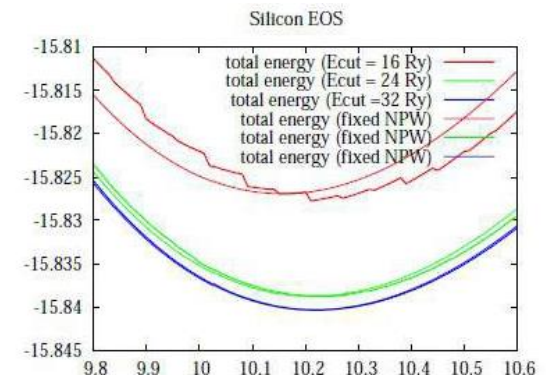


- Resulting deformations of the lattice: $\vec{h}' = (\vec{I} + \vec{\epsilon})\vec{h}$
 - ← tensor containing the lattice vectors
 - ← strain tensor

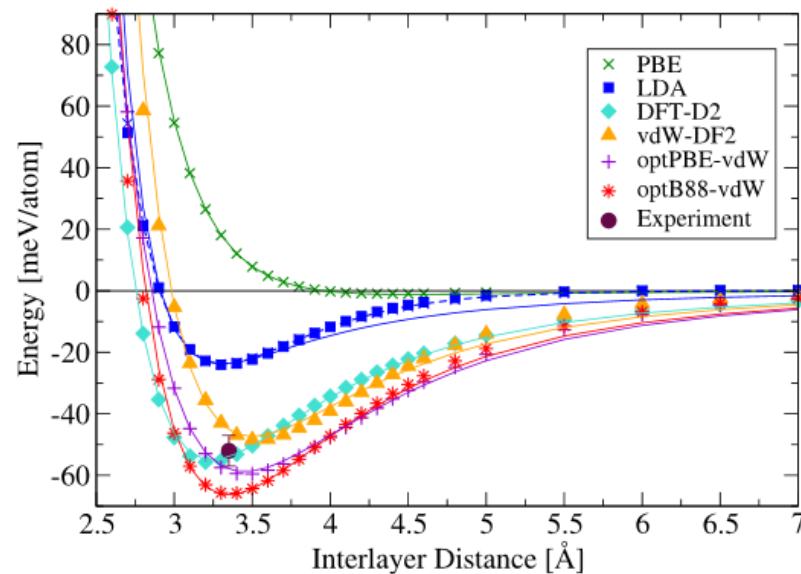
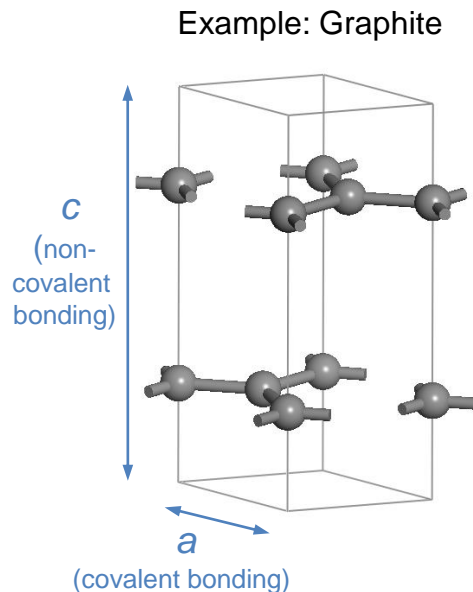
- Stress tensor from Hellmann-Feynman theorem: $\sigma_{\alpha\beta} = -\frac{1}{V} \frac{\partial E_{tot}}{\partial \varepsilon_{\alpha\beta}} = -\frac{1}{V} \sum_{\gamma} \frac{\partial E_{tot}}{\partial \vec{h}_{\alpha\gamma}} \vec{h}_{\gamma\beta}^T$
- Can be evaluated numerically, or implemented efficiently in a planewave basis (long eq.!).

- Generalized forces on lattice vectors: $F_{n\beta}^{lat} = -\frac{\partial H}{\partial a_{n\beta}} = V \sum_{\alpha} (\vec{h}^{-1})_{n\alpha} (\sigma_{\gamma\beta} - p\delta_{\gamma\beta})$

← Enthalpy $H = E_{tot} + pV$
↗ external pressure
- Use generalized forces to update components of lattice vectors during BFGS minimization
- Careful: Plane wave basis is independent of atomic positions, but NOT independent of the unit cell volume or shape
- Strong changes in unit cell can affect basis quality
- Constant number of PWs vs. constant E_{cut}

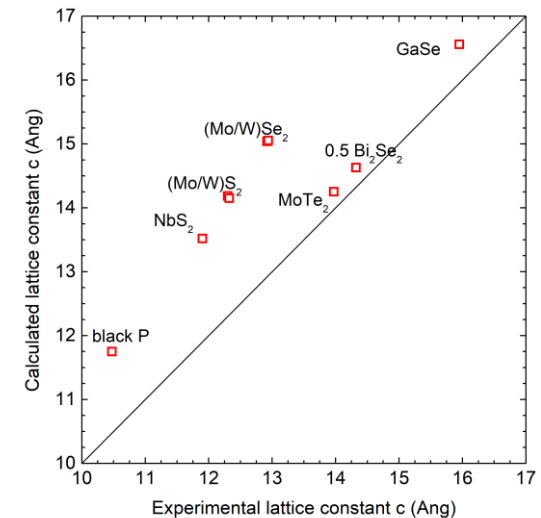
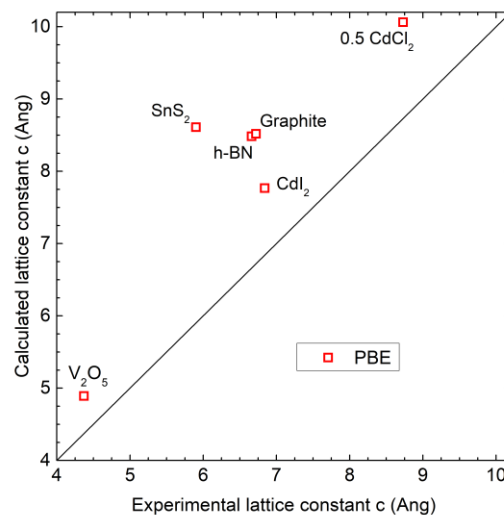
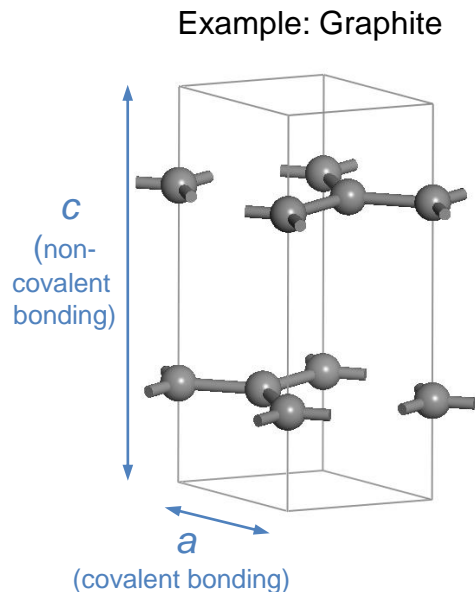


- Problem: Exchange-correlation potential of „standard“ DFT decays too strongly
→ long-range non-covalent „van-der-Waals“ bonding on surfaces or in layered materials is not well represented
- Non-covalent binding energy from „bare“ DFT usually too small



Binding energy of graphite for different functionals
J. Phys.: Cond. Mat 24, 424216 (2012)

- Problem: Exchange-correlation potential of „standard“ DFT decays too strongly
 → long-range non-covalent „van-der-Waals“ bonding on surfaces or in layered materials is not well represented
- PBE overestimates lattice constants in layered materials by ~20%



Average Error (15 layered materials)	PBE
in a (%)	1.2
in c (%)	19.7

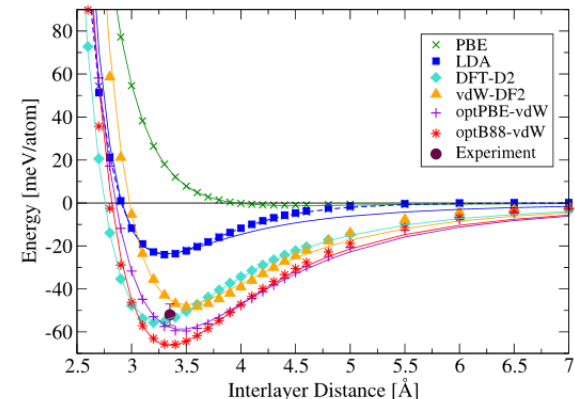
- In practice: Two common ways to include non-covalent bonding effects
- London-type corrections
 - Use semi-empirical model to correct DFT energy, forces and stresses

$$E_{tot} = E_{DFT} + \Delta E_{disp} \quad \Delta E_{disp} \propto \sum_{A,B} \frac{C_6}{R_{AB}^6} f_6(R_{AB}) + [\dots]$$

„damping“ to remove correction for small R

- Several proposals:

- DFT+D2 *J. Comp. Chem.* 27, 1787 (2006)
- DFT+D3 *J. Chem. Phys.* 132, 154104 (2010)
- DFT+TS *Phys. Rev. Lett.* 102, 073005 (2009)
Phys. Rev. Lett. 108, 236402 (2012)
- ...



- Advantages: Relatively easy and straight-forward to handle and implement
- Disadvantages: only „correction“, difficult to achieve flexibility with respect to bonding conditions

Flexibility in the DFT-D3 method

- Example: DFT-D3 $\Delta E_{disp} \propto \sum_{A,B} \frac{C_6^{AB}}{R_{AB}^6} f_6(R_{AB}) + \frac{C_8^{AB}}{R_{AB}^8} f_8(R_{AB})$
- Flexibility through pair-specific C_6^{AB} :
 - Calculate coordination number CN for atoms A and B
 - Interpolate C_6^{AB} from tabulated reference data

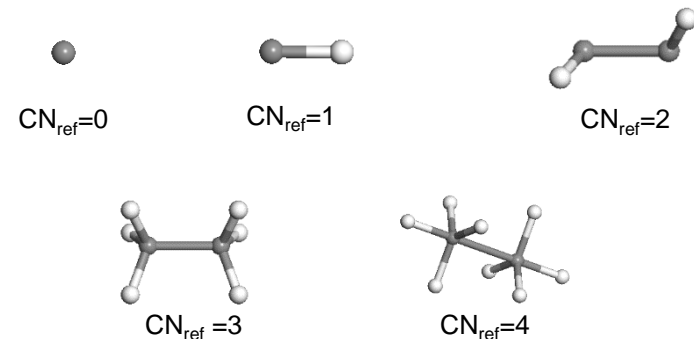
$$C_6^{AB} = \frac{1}{\sum_{i,j} L_{ij}} \left(\sum_{i,j} C_{6,ref}^{AB}(CN_i^A, CN_j^B) L_{ij} \right)$$

Sum over tabulated reference molecules for elements of A and B

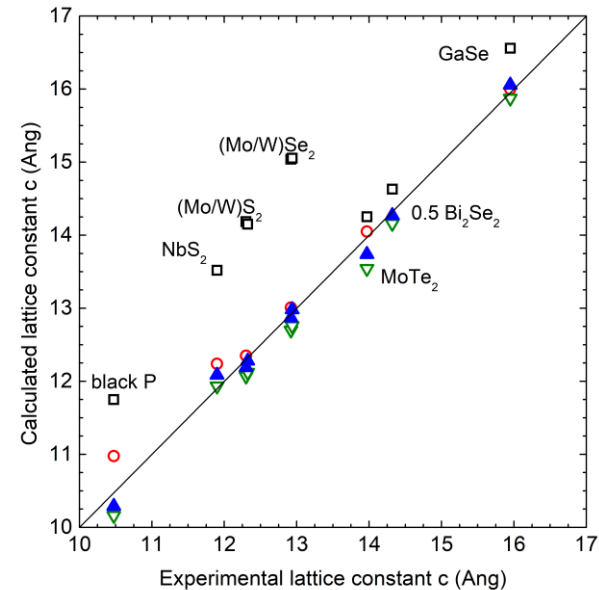
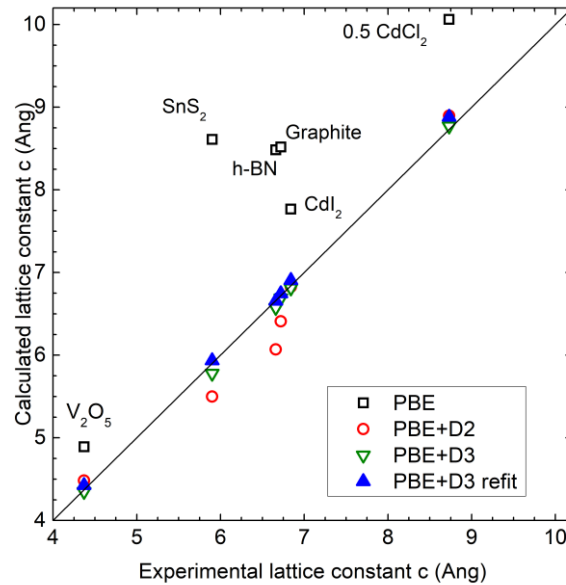
Tabulated „support points“ for interpolation

weight

$$C_8^{AB} \propto C_6^{AB}$$



- Resulting lattice constants of layered materials are significantly improved compared to PBE



Average Error (15 layered materials)	PBE	PBE-D2	PBE-D3	PBE-D3 refit
in a (%)	1.2	0.62	0.73	0.66
in c (%)	19.7	2.5	1.2	0.75

- Non-local correlation methods:

- Include non-covalent binding through a density-dependent correlation functional

$$E_{tot} = E_{DFT} + E_c^{nl}[n] \quad E_c^{nl}[n] = \iint n(r)n(r')\phi(r,r')drdr'$$

non-local „vdW kernel“ that
is tailored to complement a
specific XC functional

- Several proposals:

- vdW-DF/DF2/DF3 and variations

Phys. Rev. Lett. **92**, 246401 (2004)

- rVV10

Phys. Rev. B **87**, 041108(R) (2013)

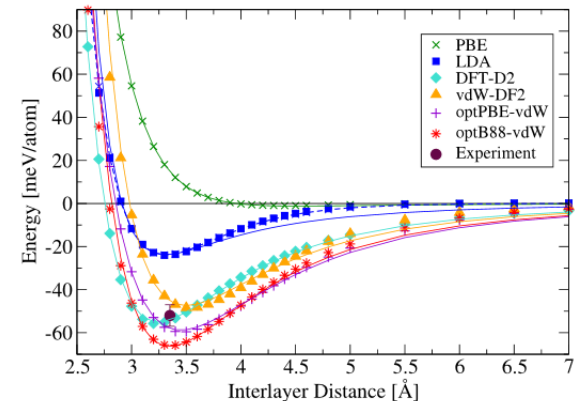
- opt88-vdW

Phys. Rev. B **82**, 081101 (2010)

- BEEF-vdW

Phys. Rev. B, **85**, 235149 (2012)

- ...



- Advantages: More consistent than London-type corrections, potentially more accurate and flexible
- Disadvantages: More difficult to develop and implement

van-der-Waals interactions should be included

- Non-local vdW functionals perform well for layered materials compared to „bare“ DFT

Method	Graphite	h-BN	TiS ₂	TiSe ₂	MoS ₂	MoSe ₂	MoTe ₂	HfTe ₂	WS ₂	WSe ₂	ME	MAE	MRE	MARE
<i>E_b</i> (binding energy between layers)														
Without dispersion:														
LDA	10 (−48)	10 (−30)	20 (7)	21 (24)	13 (−35)	14 (−29)	15 (−26)	19 (3)	13 (−37)	13 (−32)	−4	5	−20.4	27.4
TM	11 (−38)	12 (−19)	13 (−31)	14 (−19)	10 (−50)	11 (−42)	13 (−36)	13 (−28)	10 (−50)	11 (−44)	−7	7	−35.7	35.7
SCAN	7 (−59)	8 (−45)	6 (−68)	6 (−64)	6 (−73)	5 (−72)	7 (−65)	7 (−60)	6 (−72)	5 (−73)	−12	12	−65.2	65.2
PBEsol	2 (−92)	2 (−86)	7 (−62)	10 (−44)	3 (−84)	5 (−75)	8 (−62)	10 (−45)	3 (−86)	4 (−77)	−13	13	−71.4	71.4
PBE	1 (−97)	1 (−96)	1 (−93)	2 (−90)	1 (−97)	1 (−97)	1 (−94)	2 (−90)	1 (−97)	1 (−97)	−18	18	−94.8	94.8
With dispersion:														
SCAN+rVV10	20 (7)	19 (34)	18 (−3)	18 (3)	20 (−3)	19 (−1)	21 (2)	19 (0)	21 (4)	20 (−1)	1	1	4.3	5.8
PBE+rVV10L	15 (−19)	14 (−4)	19 (2)	20 (18)	19 (−6)	20 (2)	22 (4)	20 (6)	19 (−5)	20 (1)	−0	1	−0.1	6.7
vdW-DF2	20 (8)	19 (29)	19 (1)	18 (2)	19 (−6)	18 (−9)	16 (−21)	15 (−19)	19 (−5)	18 (−10)	−1	2	−3.0	10.9
vdW-DF	20 (12)	19 (35)	19 (0)	18 (2)	19 (−7)	18 (−10)	16 (−21)	15 (−18)	19 (−5)	18 (−11)	−1	2	−2.3	12.0
PBEsol+rVV10s	12 (−32)	12 (−20)	20 (7)	21 (21)	17 (−15)	17 (−11)	21 (3)	22 (19)	17 (−14)	17 (−12)	−1	3	−5.4	15.3
rev-vdW-DF2	23 (23)	21 (47)	25 (30)	24 (40)	23 (14)	22 (15)	23 (9)	22 (16)	23 (14)	22 (12)	4	4	22.0	22.0
PBE-D3(BJ)	17 (−9)	16 (9)	27 (45)	30 (72)	24 (17)	26 (34)	30 (44)	27 (46)	26 (28)	28 (38)	6	7	32.5	34.2
vdW-DF-cx	25 (36)	24 (67)	27 (43)	27 (59)	25 (21)	25 (26)	26 (25)	25 (35)	24 (21)	24 (23)	6	6	35.6	35.6
optB88-vdW	27 (47)	26 (80)	27 (45)	26 (52)	26 (28)	25 (29)	24 (16)	23 (23)	26 (29)	25 (27)	7	7	37.5	37.5
optB86b-vdW	27 (47)	26 (80)	28 (48)	28 (60)	26 (29)	26 (31)	26 (24)	25 (33)	26 (30)	26 (28)	7	7	41.0	41.0
rVV10	26 (44)	25 (72)	28 (48)	29 (65)	29 (42)	29 (50)	29 (40)	26 (40)	29 (44)	29 (48)	9	9	49.2	49.2
C09-vdW	29 (59)	28 (96)	32 (72)	33 (88)	30 (44)	29 (49)	30 (44)	30 (59)	29 (44)	29 (45)	11	11	59.9	59.9
revPBE-D3(BJ)	26 (41)	25 (71)	48 (153)	51 (196)	45 (118)	49 (152)	55 (163)	46 (147)	50 (147)	53 (167)	26	26	135.5	135.5
Reference	18.3	14.4	18.8	17.3	20.5	19.6	20.8	18.6	20.2	19.9				

Phys. Rev. Mat. 3, 063602 (2019)

- Also good performance for lattice constants (also for covalently bonded crystals!)
- Computationally efficient, only slightly slower than LDA or GGAs
- Should always be used for systems with possible non-covalent bonding