An Introduction to Density-Functional Theory for Experimentalists

Erin R. Johnson

http://schooner.chem.dal.ca

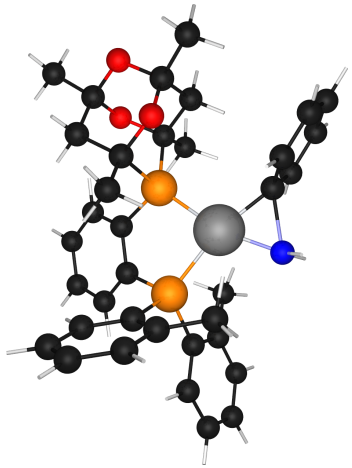
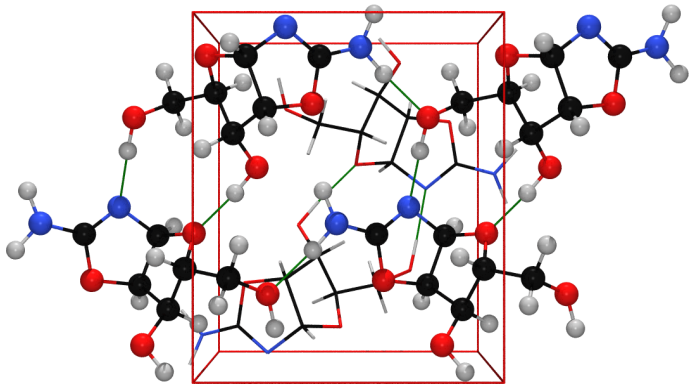
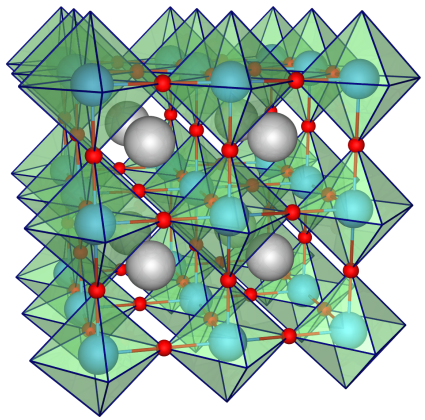
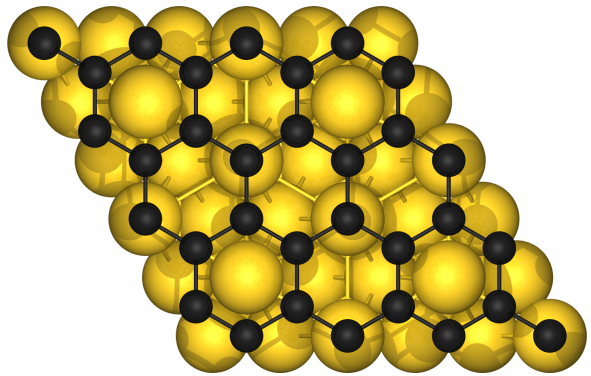
Partnership between theory and experiment

Simulating chemical processes computationally can help interpret and inform experiment.

What can theory predict?

The aim of quantum chemistry is to accurately predict molecular and materials properties from first principles.

- molecular or crystal geometries
- thermodynamic quantities
- mechanisms and chemical kinetics
- spectroscopic quantities
- mechanical properties

What details do you need to include?

When reading a theory section, you should be able to find enough detail for the calculation to be reproducible.

This should include:

- Geometry details for the system under study
- The choice of density functional (and dispersion correction)
- Basis set or plane-wave/pseudopotential details
- Any other specific options

Crystals

Geometries of known crystals are available through three databases:

Crystallography Open Database
`http://crystallography.net/doc`

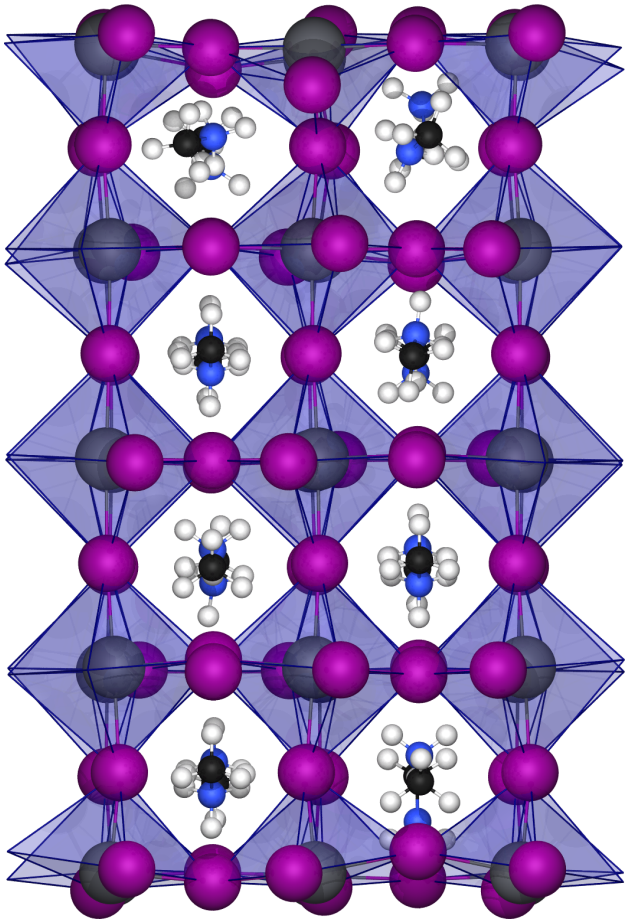
Cambridge Structure Database
`http://ccdc.cam.ac.uk`

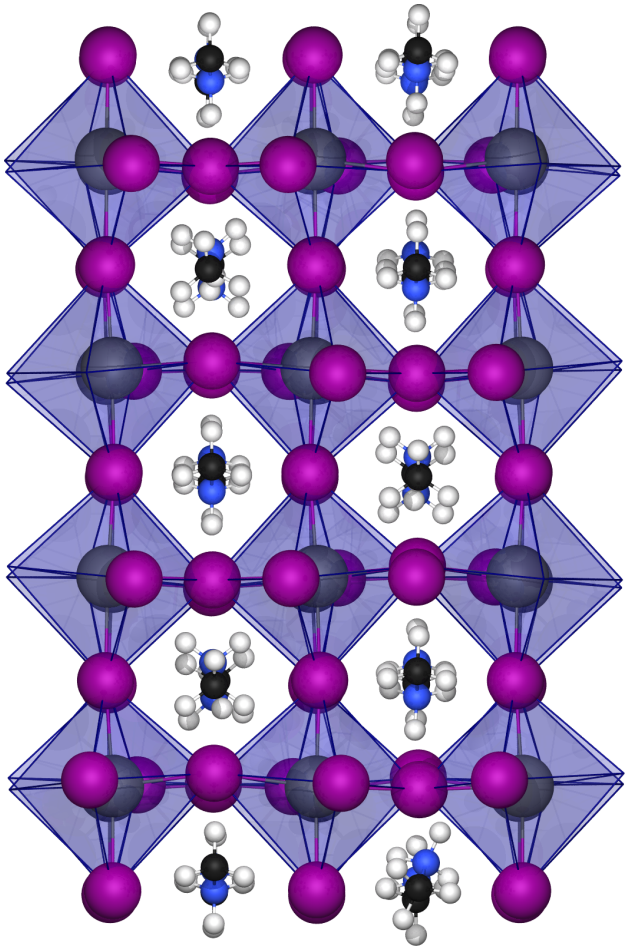
Inorganic Crystal Structure Database
`http://icsd.products.fiz-karlsruhe.de`

Report the structure code/origin of each cif file.

Surfaces

Must specify surface termination and any reconstruction – include geometry in the SI.



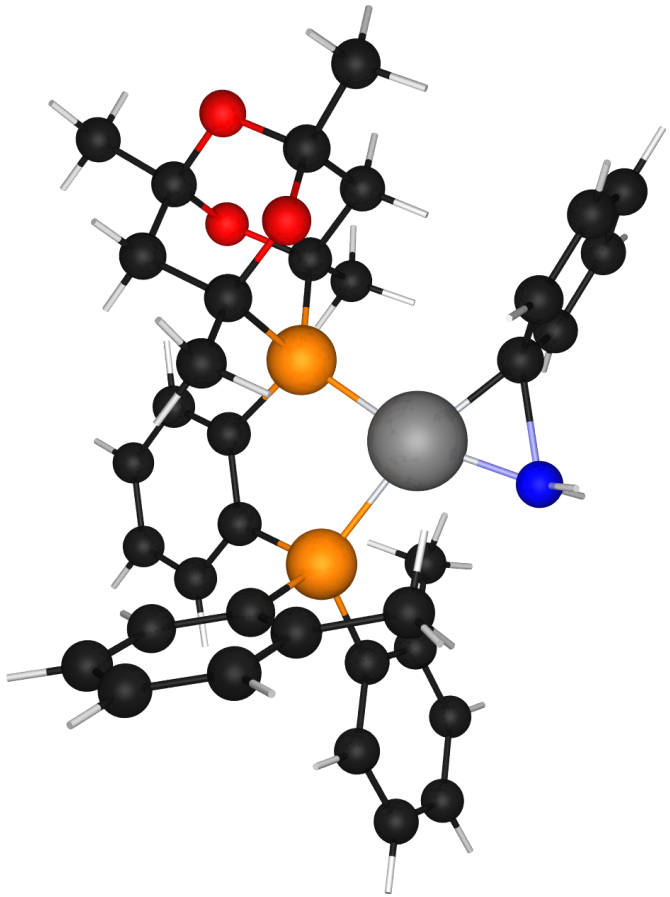


Molecules

Optimized xyz coordinates for all structures should appear in the SI.

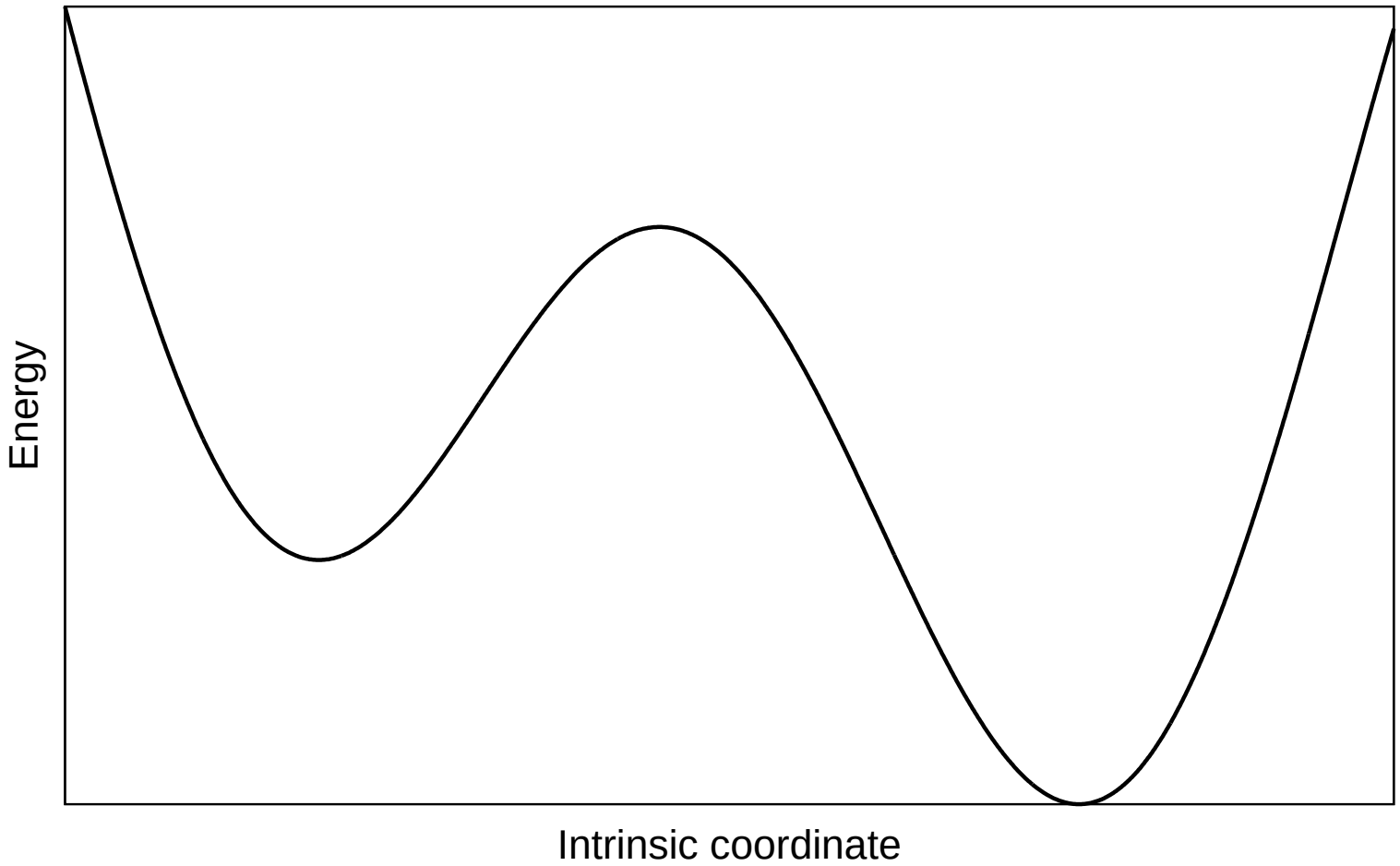
- Also include charge/multiplicity when it is not obvious.

```
Ni 0.342223 0.007233 -1.836928
P -1.222833 -0.358973 -0.504922
P 1.955490 0.028431 -0.512822
C -2.187686 1.149487 0.155336
C -0.249508 -1.048898 0.930789
C -2.804241 -1.416203 -0.614098
C 3.612158 -0.783027 -0.608310
C 1.162462 -0.892194 0.904904
C 2.344338 1.682090 0.208011
C -3.007420 0.802391 1.399961
O -3.098163 1.529544 -0.903170
C -1.270429 2.338465 0.373163
C -0.824919 -1.777121 1.984884
C -3.660874 -0.725084 -1.686323
O -3.567367 -1.450836 0.616926
```

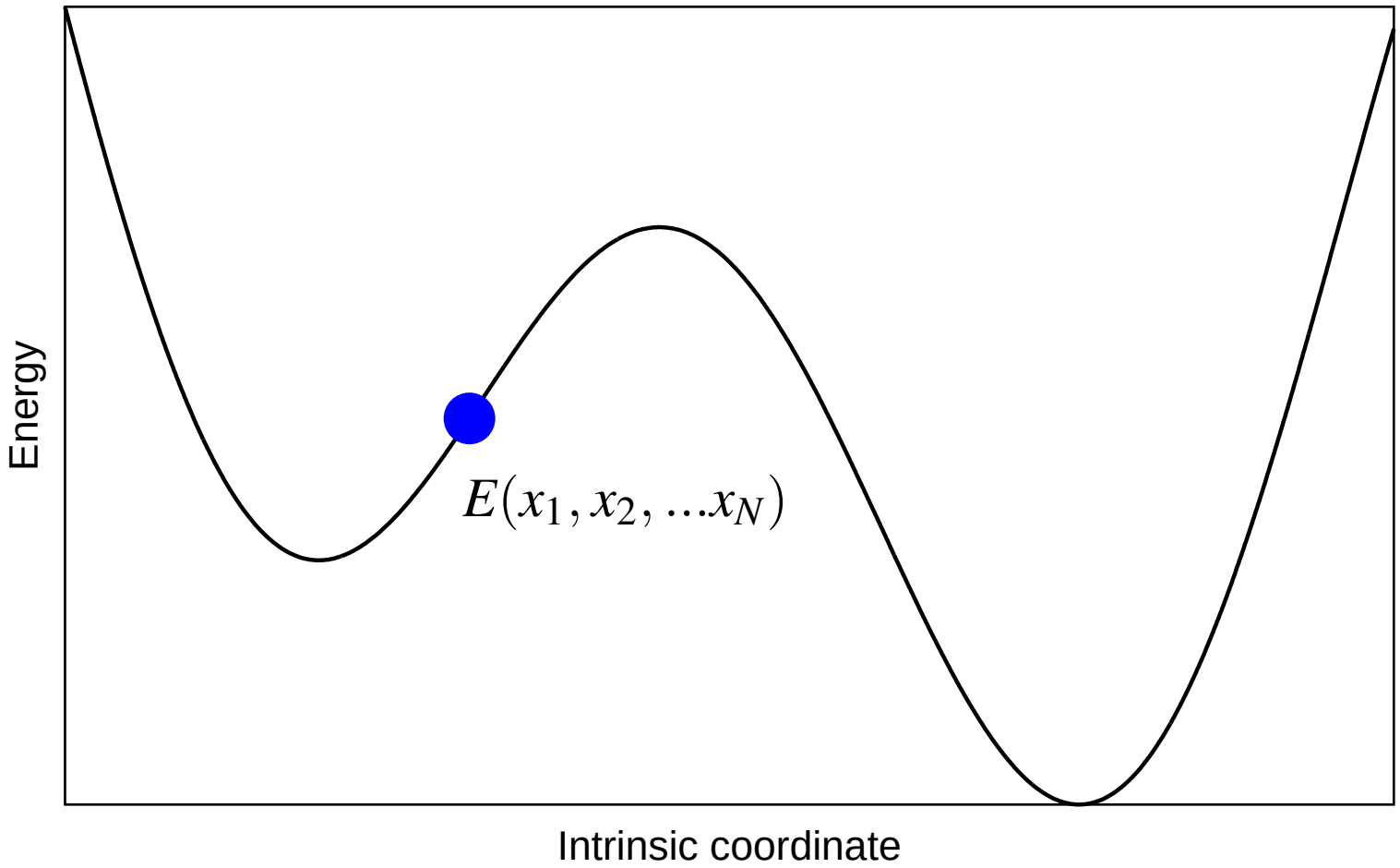


What types of calculations are possible?

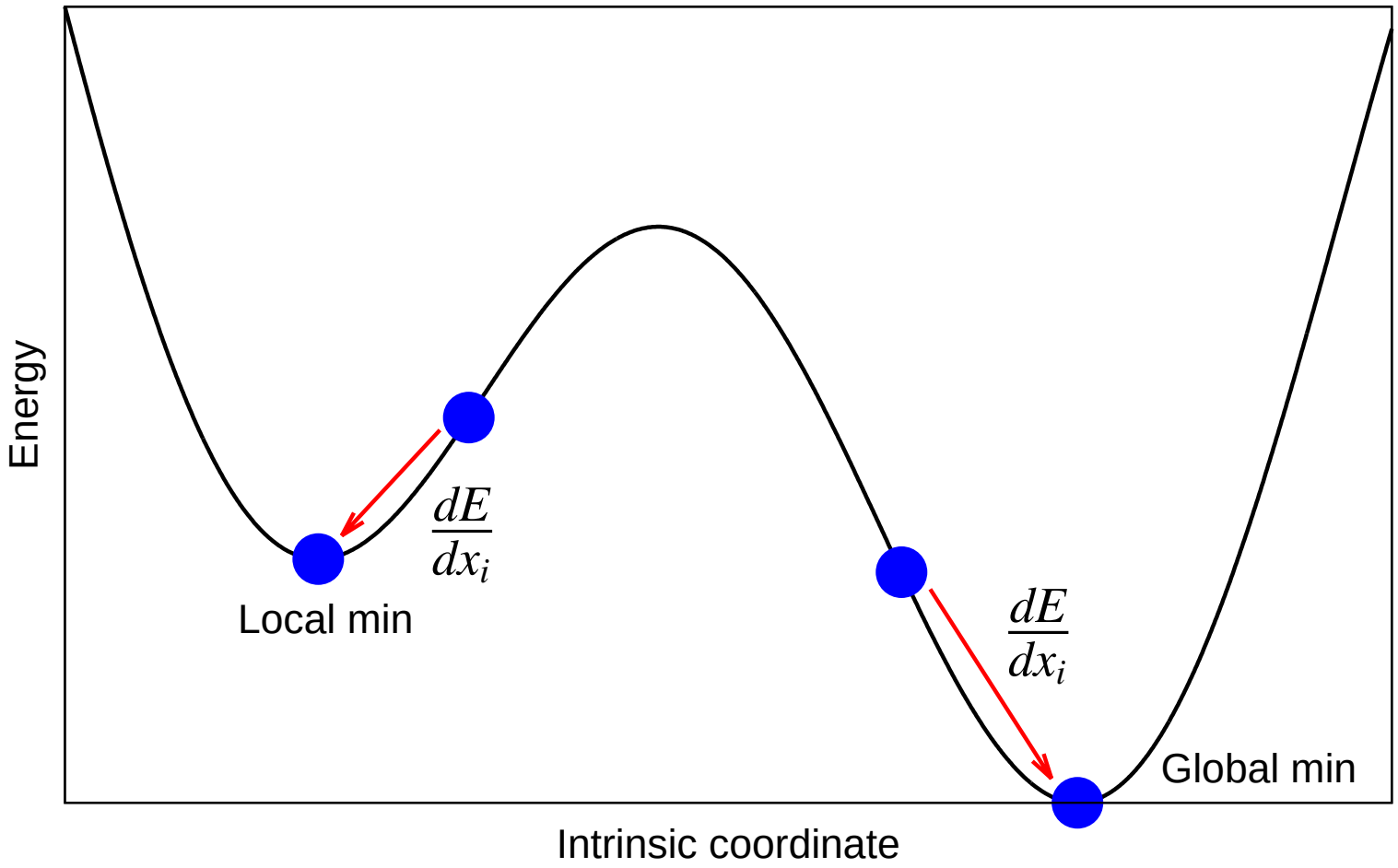
Consider a potential energy surface (PES):



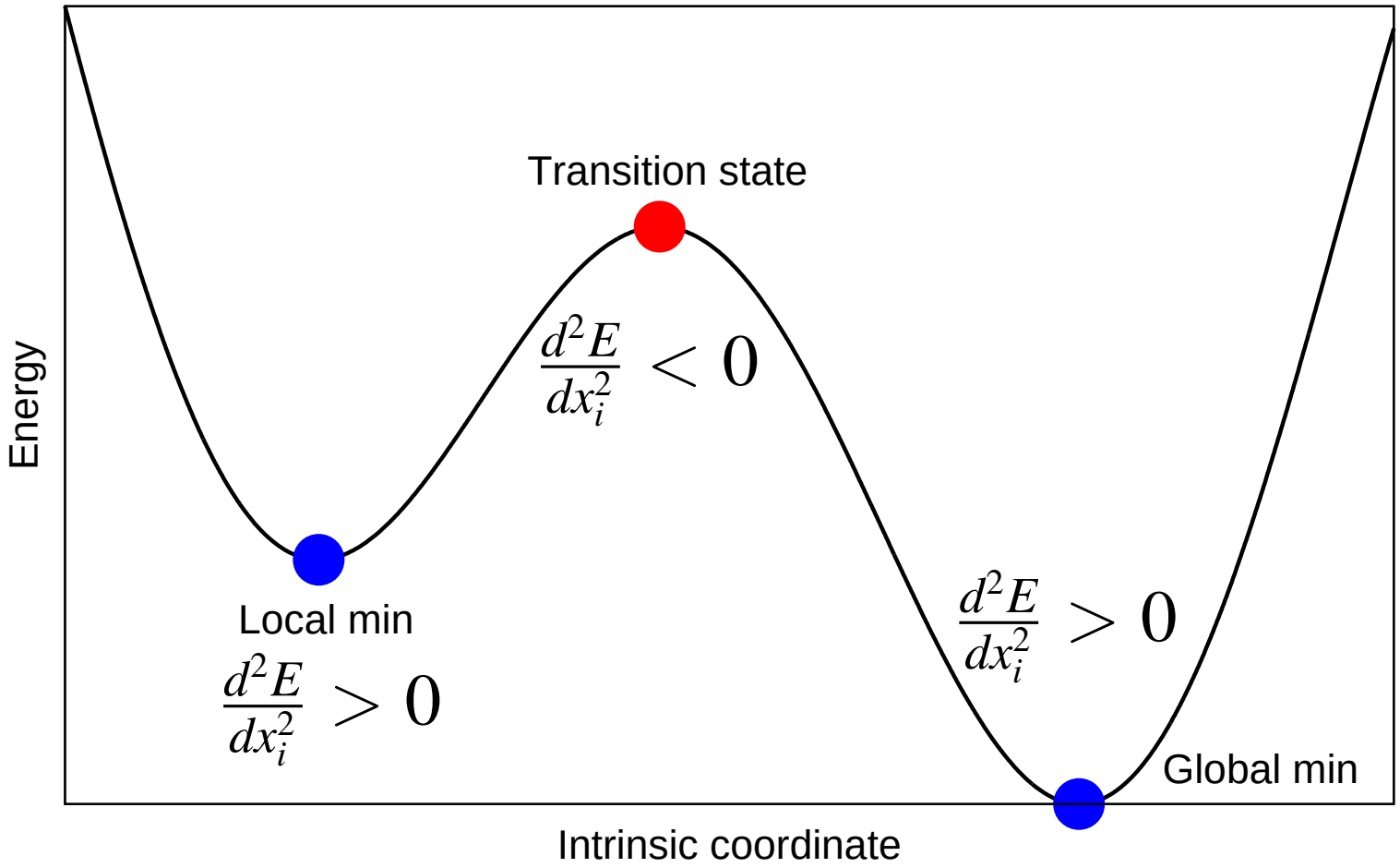
Calculation types — single-point energy



Calculation types — geometry optimization



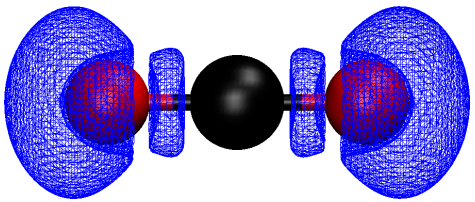
Calculation types — frequencies

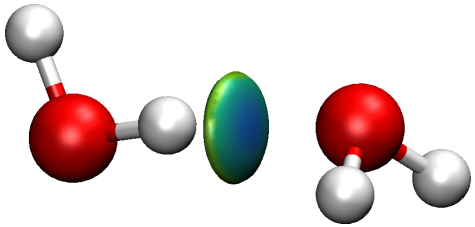


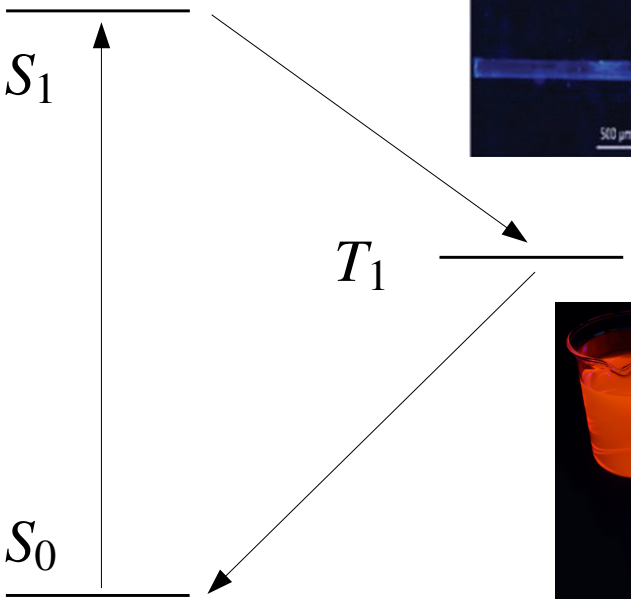
Other calculation types

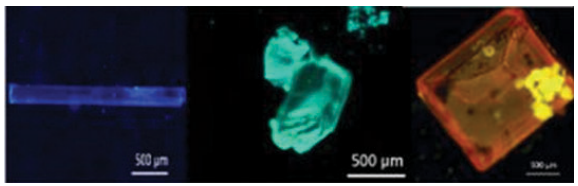
Detailed analysis of a particular property can also be done as part of a single-point energy calculation.

- Charge analysis
- Energy decomposition
- ELF and NCI plots
- Band structures
- Density of states
- Excitation energies
- NMR shieldings



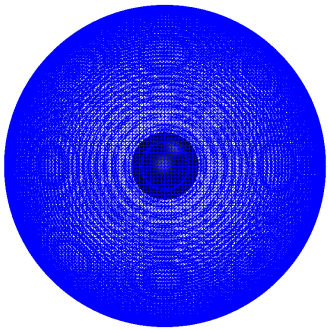
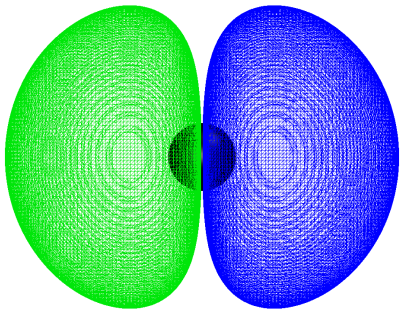
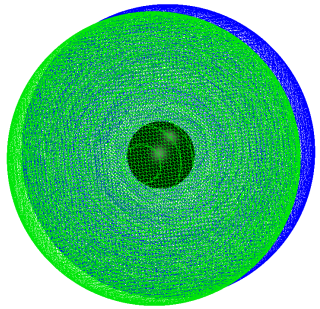
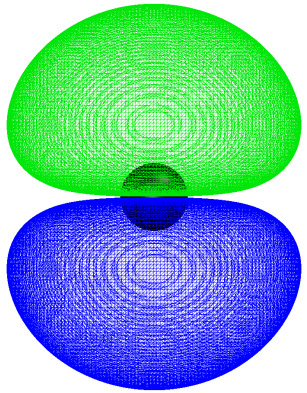






Atomic orbitals

Take linear combinations of atomic orbitals to form molecular orbitals:

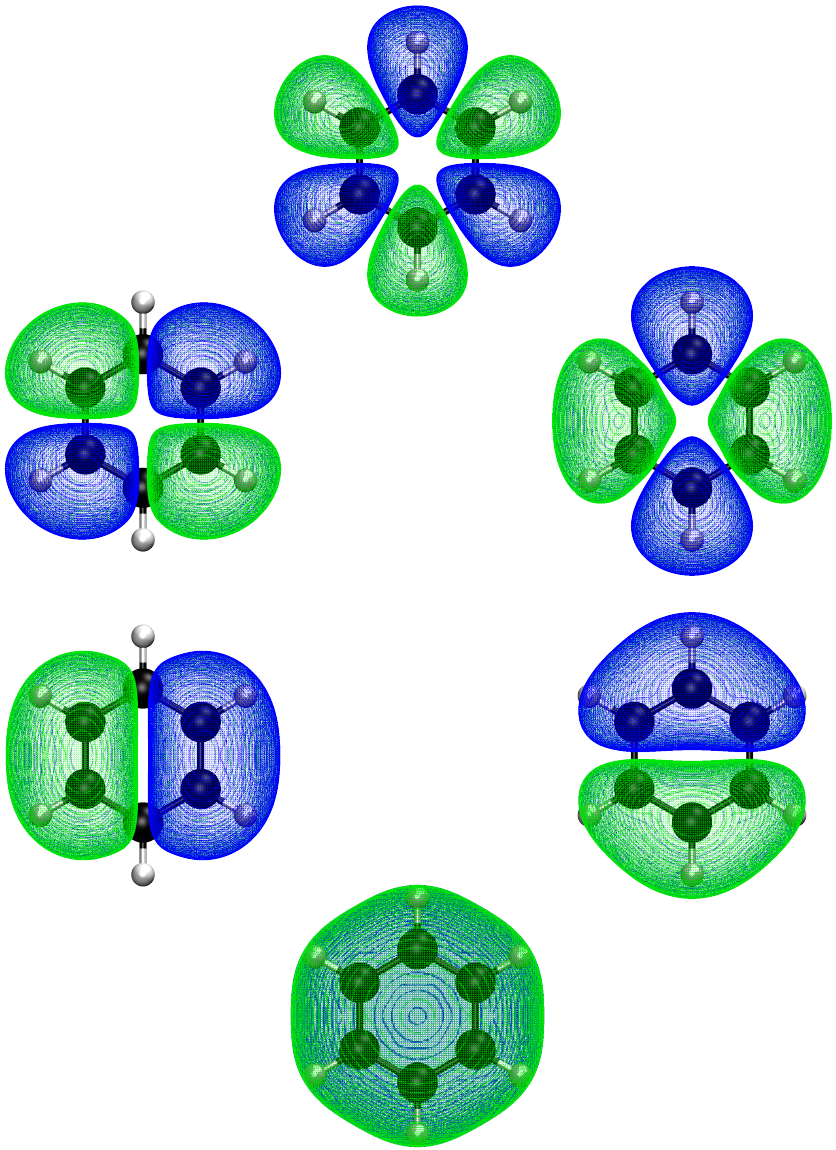





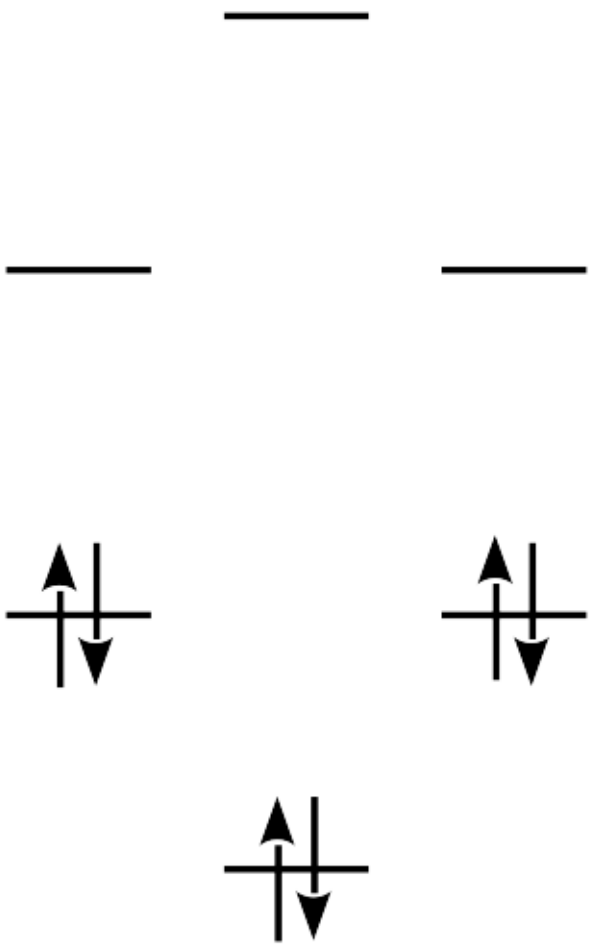
$$MO = c_1 AO_1 + c_2 AO_2 + \ldots + c_n AO_n = \sum_{i=0}^{n} c_i AO_i$$

$$\phi = c_1 \chi_1 + c_2 \chi_2 + \ldots + c_n \chi_n = \sum_{i=0}^{n} c_i \chi_i$$

Then fill the MOs with electrons.

Molecular orbitals





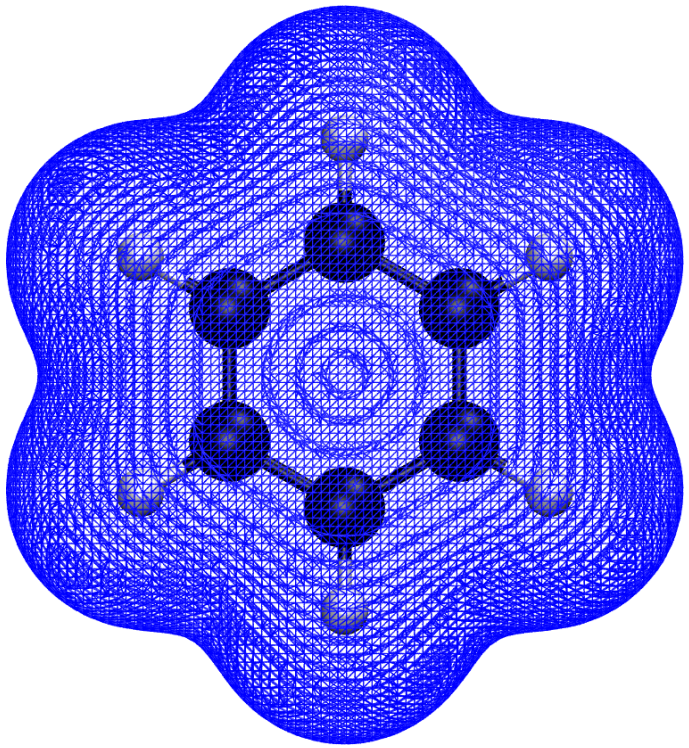
Electron density and energy

The occupied MOs generate the 3D electron density, ρ.

$$\rho(x,y,z) = \sum_i \phi_i^2(x,y,z)$$

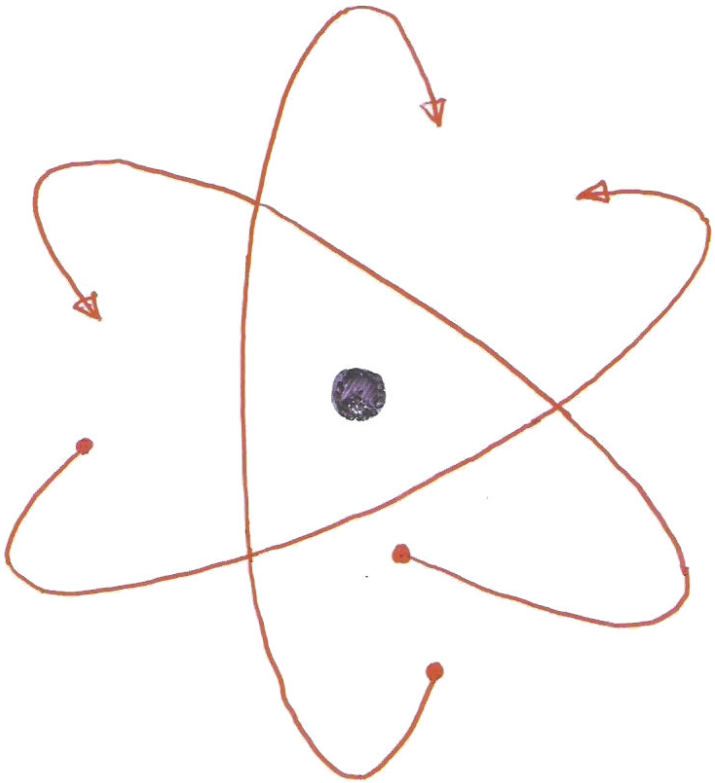
The energy is written as a functional of the density:

$$E = E[\rho(x,y,z)]$$

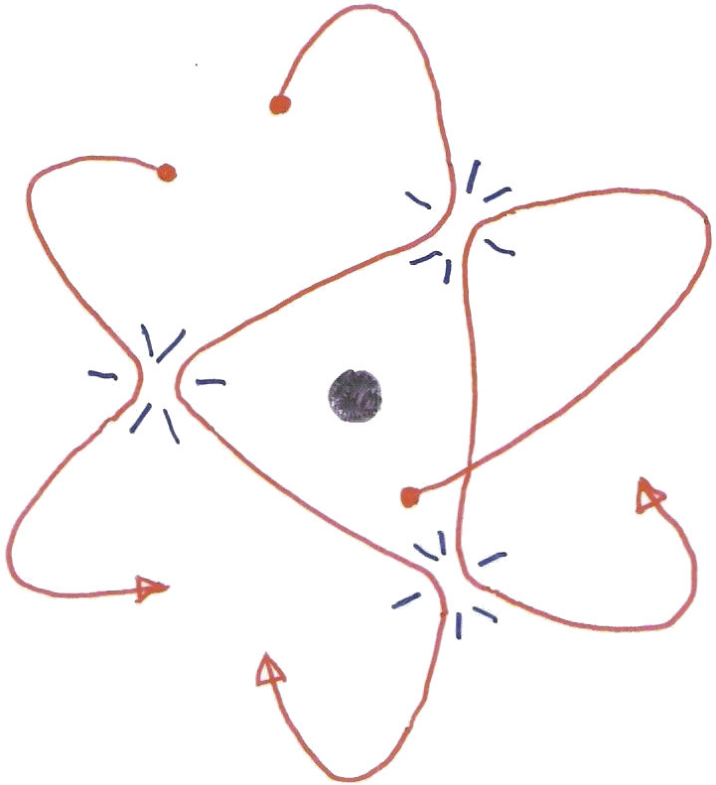


What is the challenge?

Electrons cannot be treated as classical, independent particles.



vs.



Density-functional theory (DFT)

The DFT energy is written as a sum of terms:

$$E = T_0(\rho) + V(\rho) + J_0(\rho) + E_{\mathrm{XC}}(\rho)$$

- T_0 is the kinetic energy of the electrons
- V is the electron-nuclear potential energy
- J_0 is the classical electron-electron repulsion energy
- E_{XC} is the exchange-correlation energy

E_{XC} is the difference between the classical and quantum-mechanical electron-electron interactions.

Density-functional approximations (DFAs)

The DFT energy is written as a sum of terms:

$$E = T_0(\rho) + V(\rho) + J_0(\rho) + E_{\mathrm{XC}}(\rho)$$

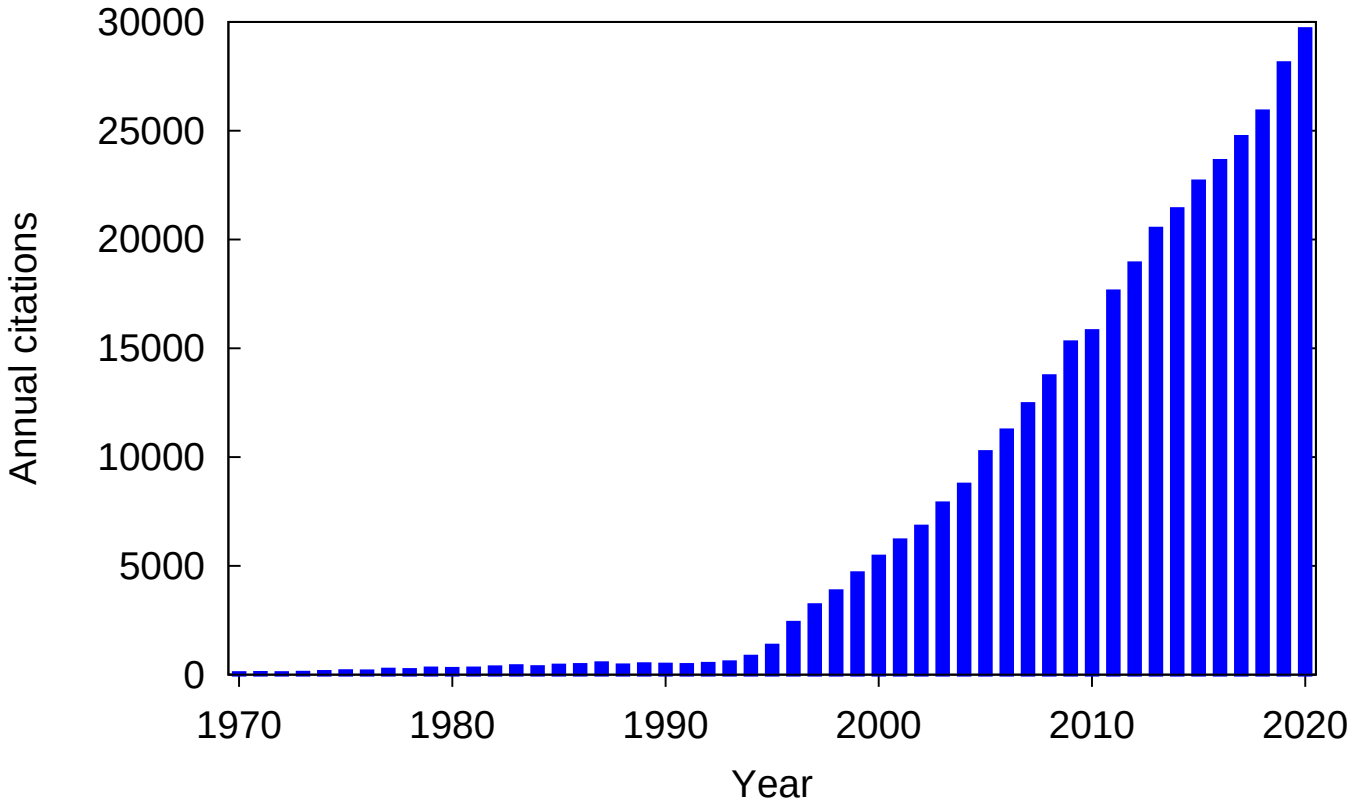
✓ ✓ ✓ ✗

The forms of T_0, V, and J_0 are known and straightforward to compute.

The exact form of E_{XC} is unknown and there is no systematic route to obtain it – very many DFAs have been proposed over 4 decades.

Density-functional approximations (DFAs)

While DFT was formulated in 1965, the most popular DFAs were developed between 1986-1996.



Walter Kohn

Axel Becke

Their use is increasing with advances in computer technologies.

Hierarchy of DFAs

In order of roughly increasing accuracy and computational time:

- The local spin-density approximation (LSDA) — $E(\rho)$
- Generalized gradient approximations (GGAs) — $E(\rho, \nabla\rho)$
  - common examples: BLYP, PW91, PBE, PBEsol, B86bPBE
- meta-GGAs — $E(\rho, \nabla\rho, \nabla^2\rho, \tau)$
  - common examples: BR, M06-L, SCAN
- Global hybrids — global mixing of DFAs with Hartree-Fock (HF)
  - common examples: B3LYP, PBE0, M06-2X
- Range-separated hybrids — variable mixing of DFAs with HF
  - common examples: LC-BLYP, LC-ωPBE, ωB97X

Density-functional thermochemistry

Mean absolute errors, in kcal/mol, for 222 heats of formation:

Method type	*MAE*
Hartree-Fock	*225*
LSDA	*120*
(meta-)GGAs	*$\sim$ 10*
Hybrids	*$\sim$ 2-5*

Hybrid functionals are quite accurate for intramolecular bonding.

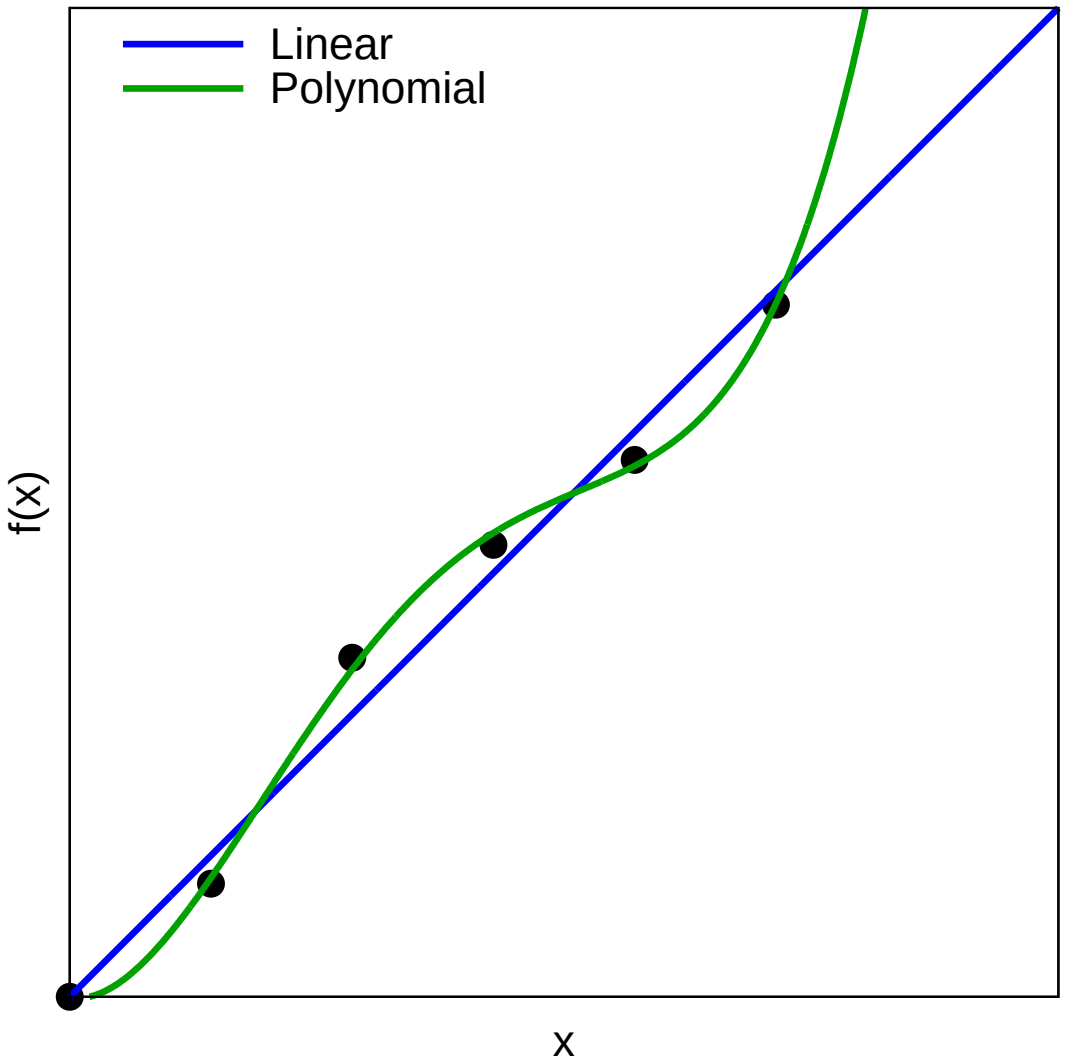
Physics vs. parameterization

Two schools of thought in functional development:

- Physical constraints
- Empirical parameters

"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk."

John von Neumann

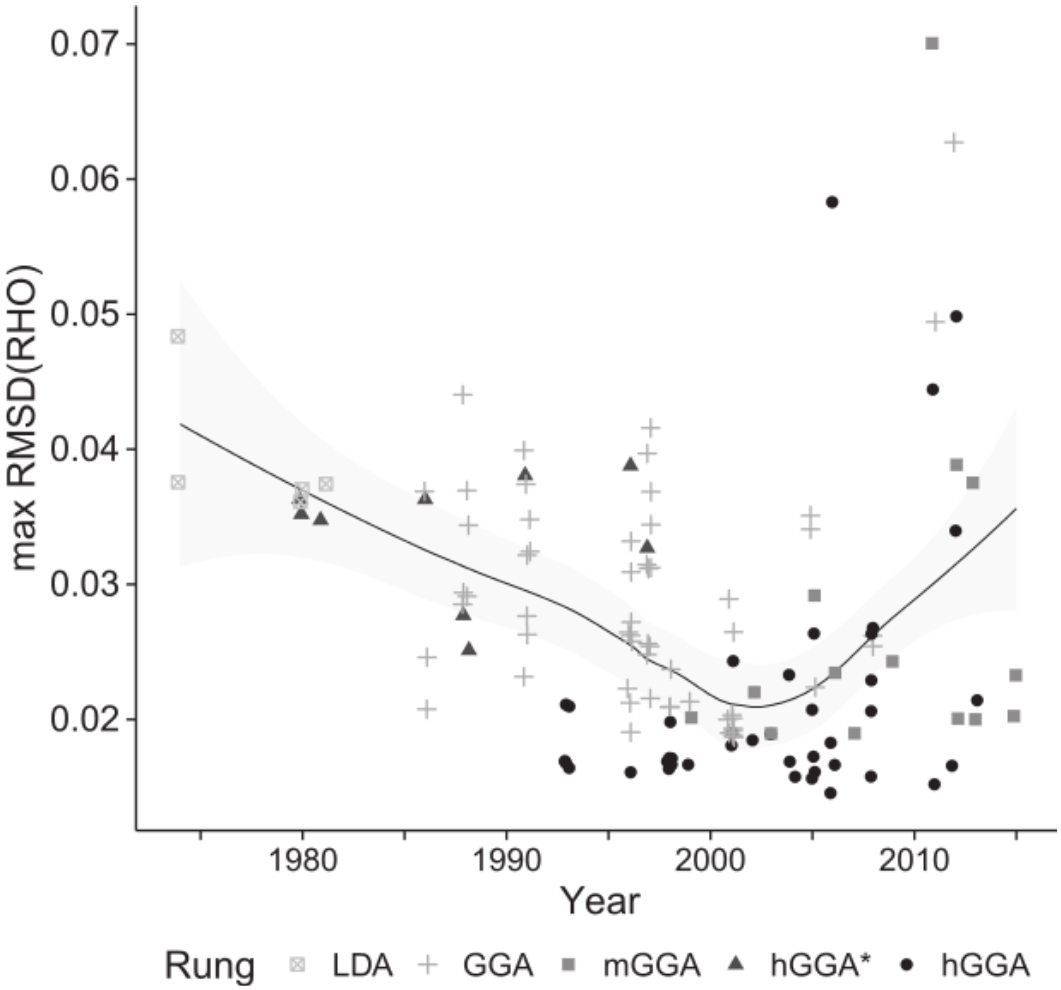


Physics vs. parameterization

Two schools of thought in functional development:

- Physical constraints
- Empirical parameters

Too many parameters can lead to over-fitting and errors outside the fit set.



Medvedev *et al.* Science **355**, 49-52 (2017)

What type of DFA should I use when?

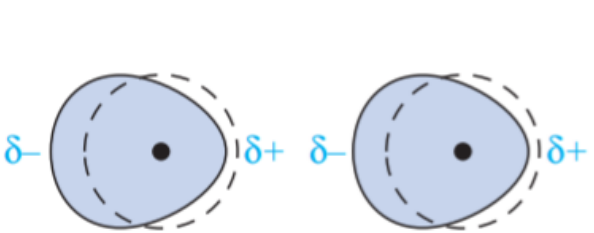
Solid-state: use GGAs, like PBE or B86bPBE, since hybrids are too expensive.

Molecular: use global hybrids like B3LYP or PBE0, unless you have:

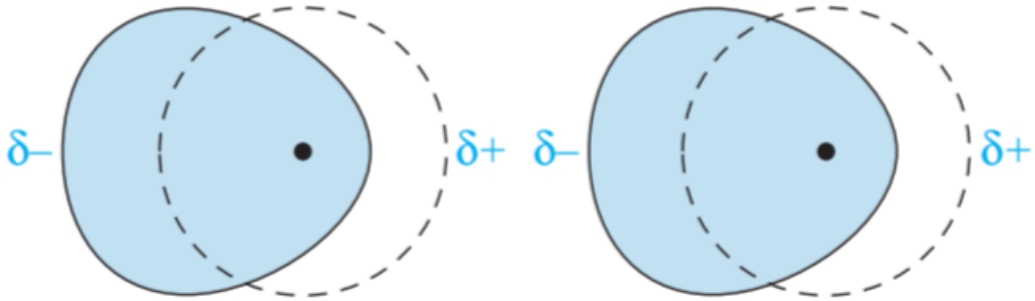
- Delocalization/self-interaction error:
  - Examples include charge-transfer complexes, halogen bonding, H-atom transfer transition states, etc.
  - use range-separated hybrids like LC-ωPBE, ωB97X
- Multi-reference systems/strong correlation error:
  - Examples include open-shell singlet biradicals ($^1O_2$), some transition-metal compounds (Cr_2)
  - Use correlated-wavefunction theory, not DFT, if possible
  - Otherwise, use GGAs, like PBE or B86bPBE

Dispersion corrections

London dispersion is responsible for condensation of non-polar species, like noble gases and hydrocarbons.



Neon (10 electrons)

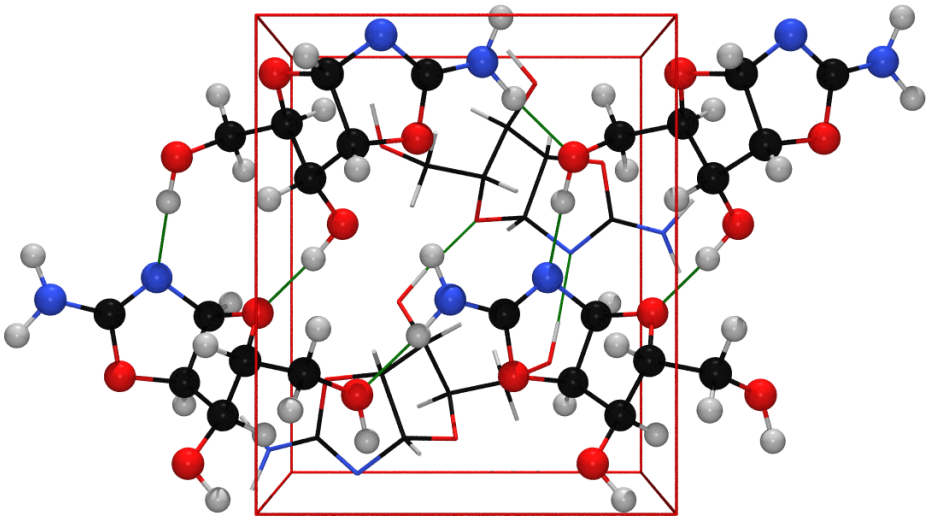


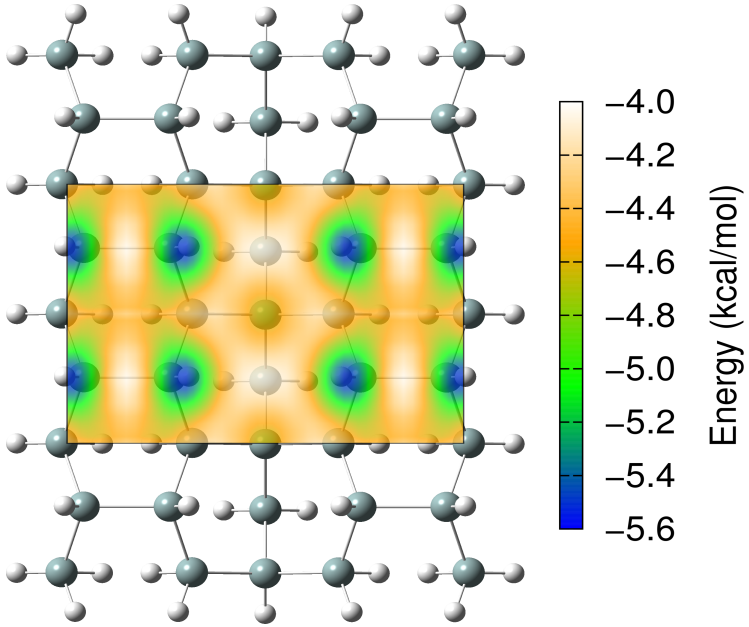
Xenon (54 electrons)

The strength of dispersion interactions increases with polarizability, explaining boiling-point trends.

This long-range non-local interaction is not captured by most density functionals.

Dispersion interactions



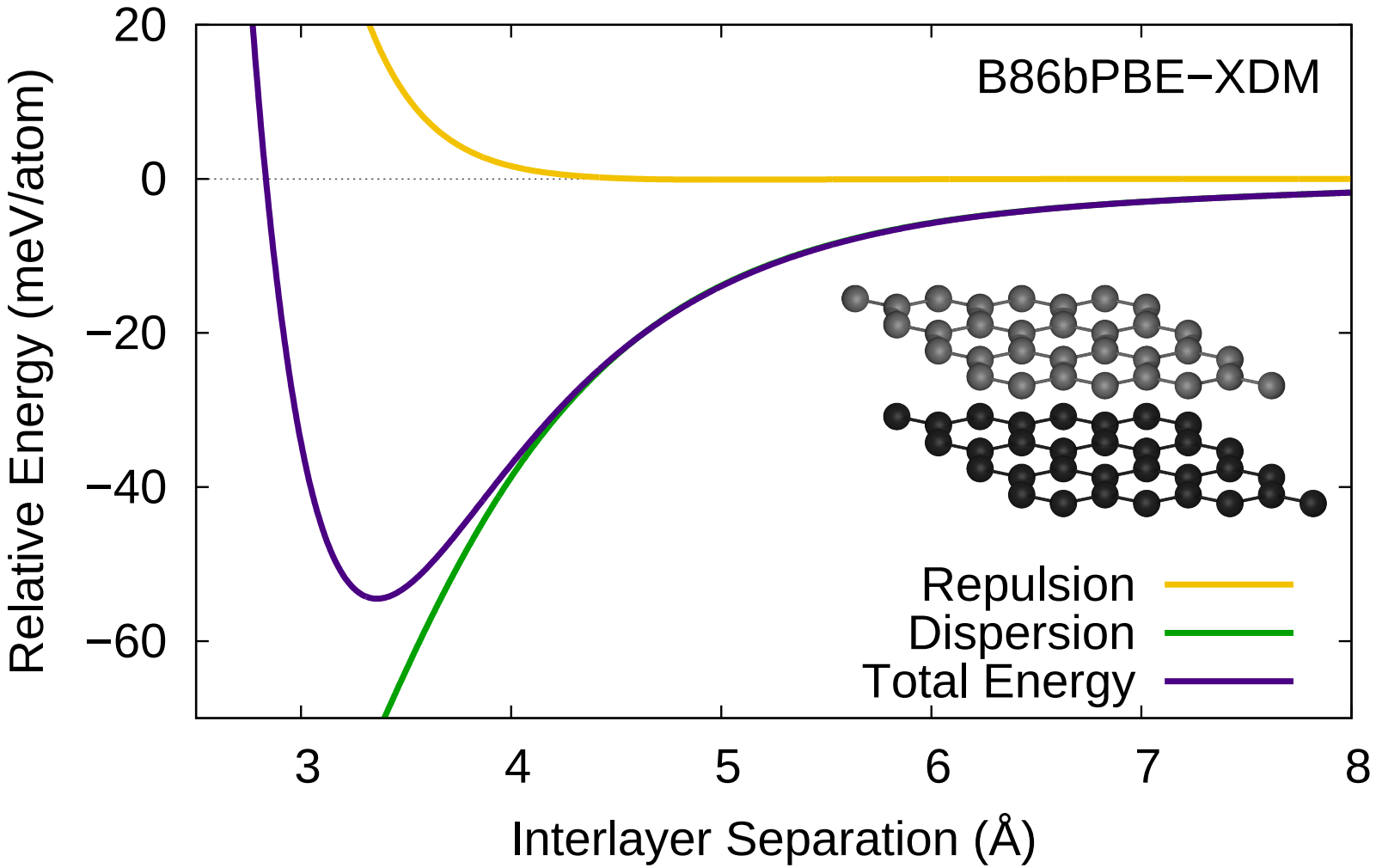


- Biomolecular structure
- Self-assembly
- Layered materials

- Surface adsorption
- Phase transitions
- Crystal packing

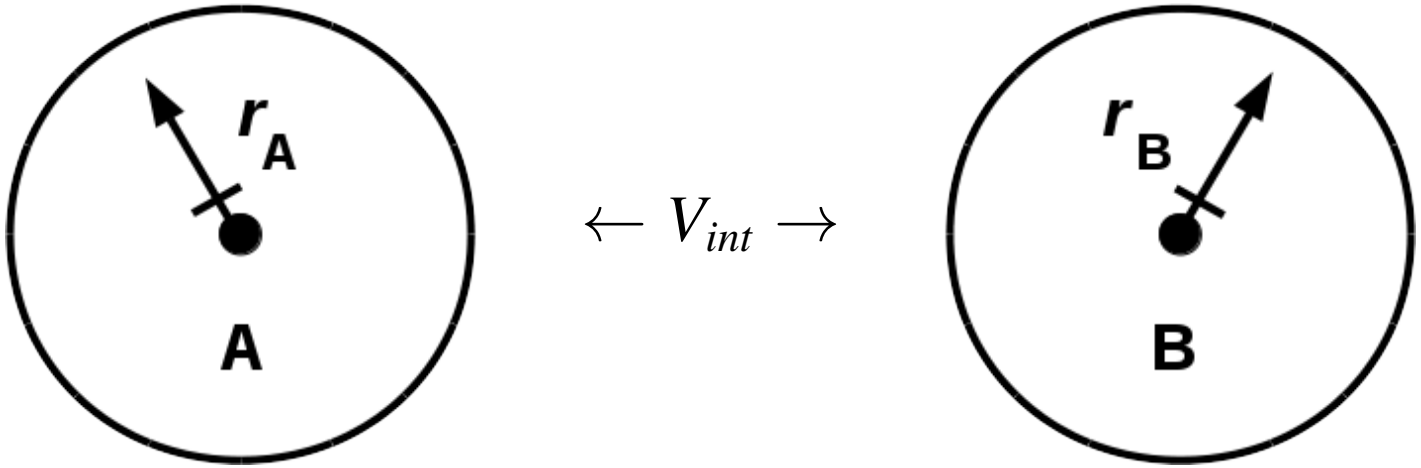
Dispersion interactions: graphite

Explicit dispersion-energy terms need to be added to DFT methods.



The dispersion energy

Dispersion arises from interaction of instantaneous dipoles (and higher-order multipoles) in the electron density distribution.



The dispersion energy can be written as a sum over all atom pairs:

$$E_{\text{disp}} = -\sum_{i<j} \left(\frac{C_6}{R_{ij}^6} + \frac{C_8}{R_{ij}^8} + \dots \right)$$

Hierarchy of dispersion corrections

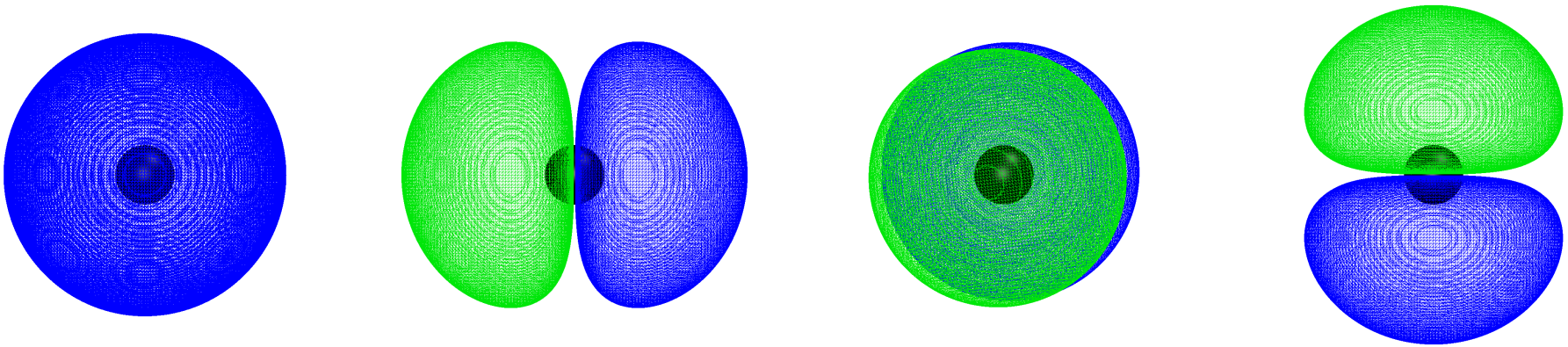
In order of roughly increasing accuracy and computational time:

- D2, TS — C_6's are fixed or weakly dependent on environment
- D3, D3(BJ) — dispersion coefficients depend on coordination
- XDM, MBD — highly dependent on environment
- vdW-DF, rVV10 — explicitly non-local XC functional

Choosing any is better than nothing — dispersion corrections are physically important and should not be viewed as optional.

Atom-centered basis sets

Take linear combinations of atomic orbitals to form molecular orbitals:



$$\phi = c_1\chi_1 + c_2\chi_2 + \ldots + c_n\chi_n = \sum_{i=0}^{n} c_i\chi_i$$

However, using hydrogen-like AOs is not practical for computing the electron-electron repulsion integrals.

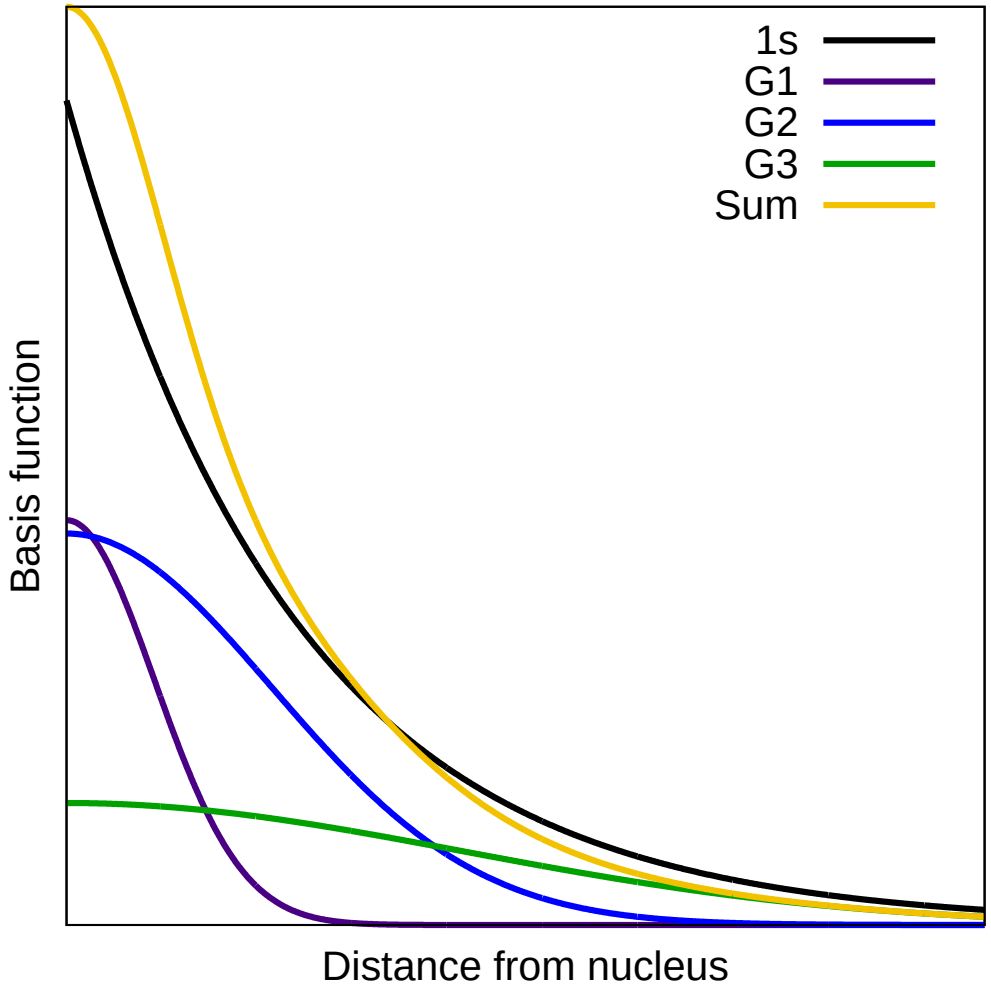
Gaussian basis sets

Use a sum of Gaussians to represent each AO:

$$\chi = Nx^i y^j z^k e^{-ar^2}$$

$$\begin{aligned} i+j+k &= 0 \quad \rightarrow \quad s \text{ orbital} \\ i+j+k &= 1 \quad \rightarrow \quad p \text{ orbital} \\ i+j+k &= 2 \quad \rightarrow \quad d \text{ orbital} \\ i+j+k &= 3 \quad \rightarrow \quad f \text{ orbital} \end{aligned}$$

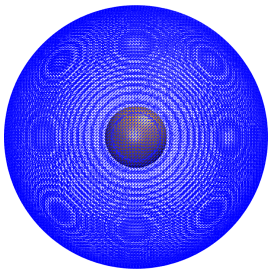
Using 2 or more sets of Gaussians per orbital allows them to expand or contract.



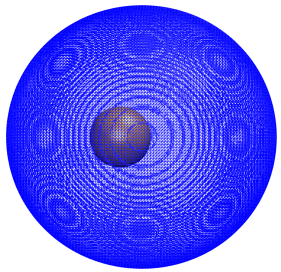
Gaussian basis sets — building flexibility

Polarization functions (*):

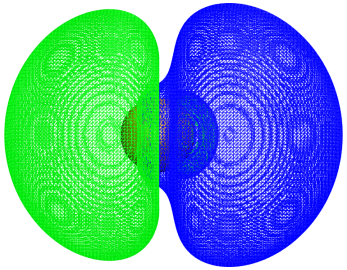
- include higher angular-momentum functions
- add p functions to s-block and d functions to p-block elements
- allows polarization of the density along bonds and for lone pairs



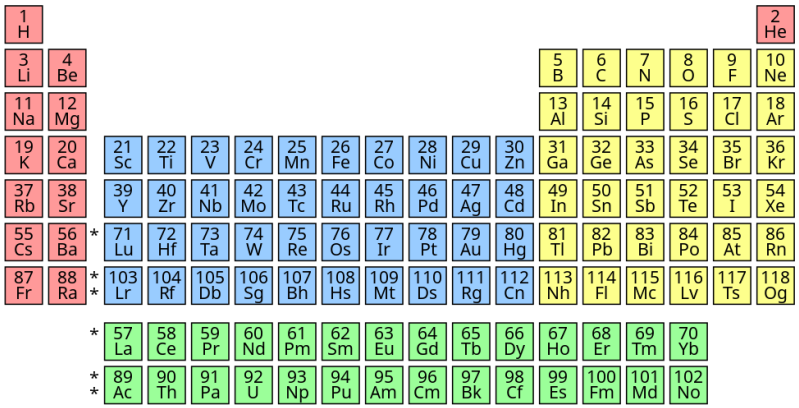
Spherical density



Polarized density



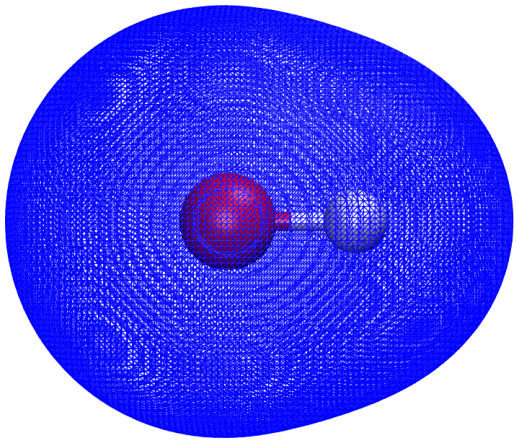
Density difference



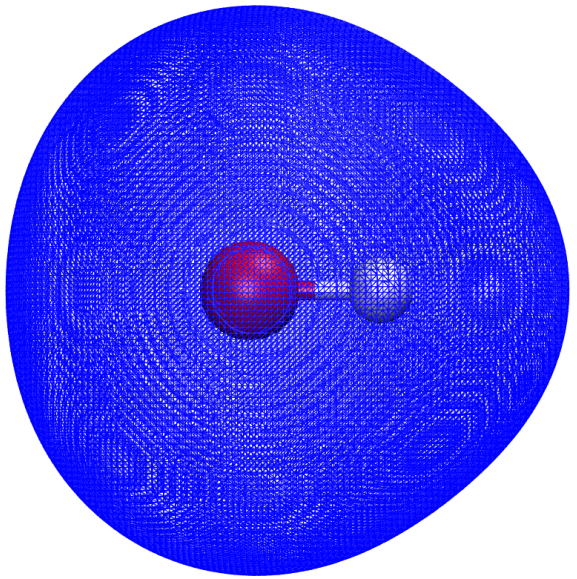
Gaussian basis sets — building flexibility

Diffuse functions (+):

- add broad Gaussians with small exponents
- important for anions, which have large atomic radii
- important for intermolecular interactions, like H-bonding



Neutral



Anion

Which basis sets should I use?

Geometry optimization and frequency calculations:

- time-consuming, require many energy evaluations
- fairly insensitive to basis set
- use small basis sets: 6-31G*, cc-pVDZ, def-SVP
- for anions and very electronegative atoms, add diffuse functions: 6-31+G*, aug-cc-pVDZ, def-SVPD

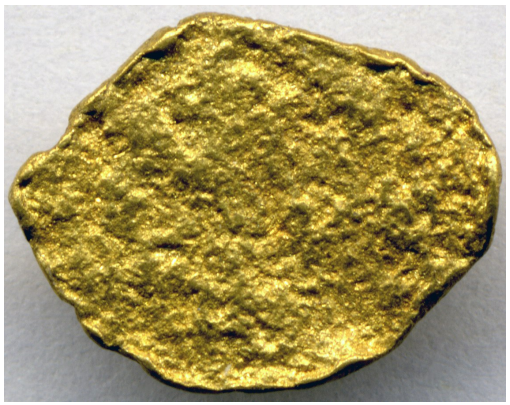
Single-point energy calculations:

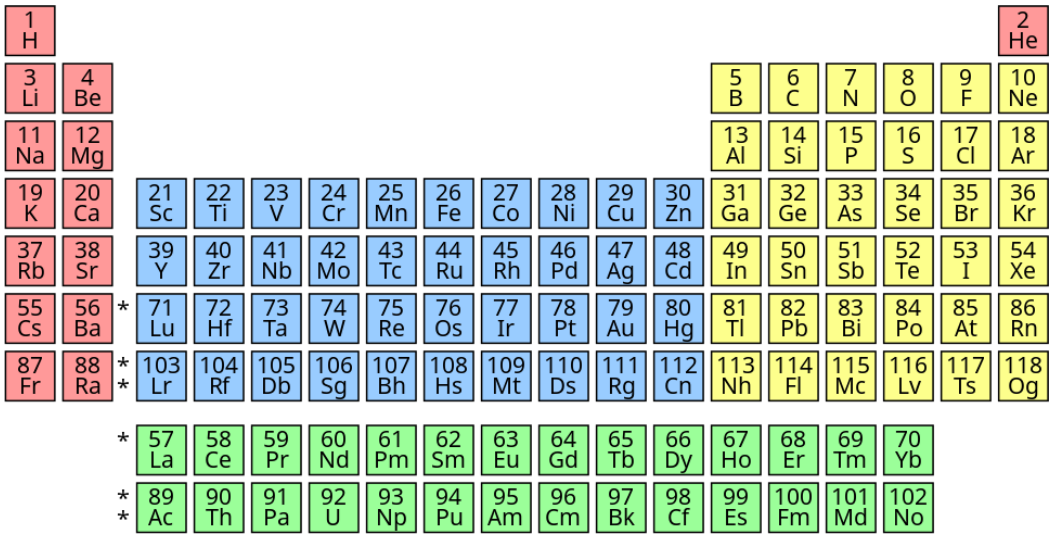
- fairly quick, require only one energy evaluation
- quite sensitive to basis set
- use large basis sets: 6-311+G(2d,2p), aug-cc-pVTZ, def-TVZPD

Effective core potentials (ECPs)

For heavy elements, core electrons are replaced by an ECP to

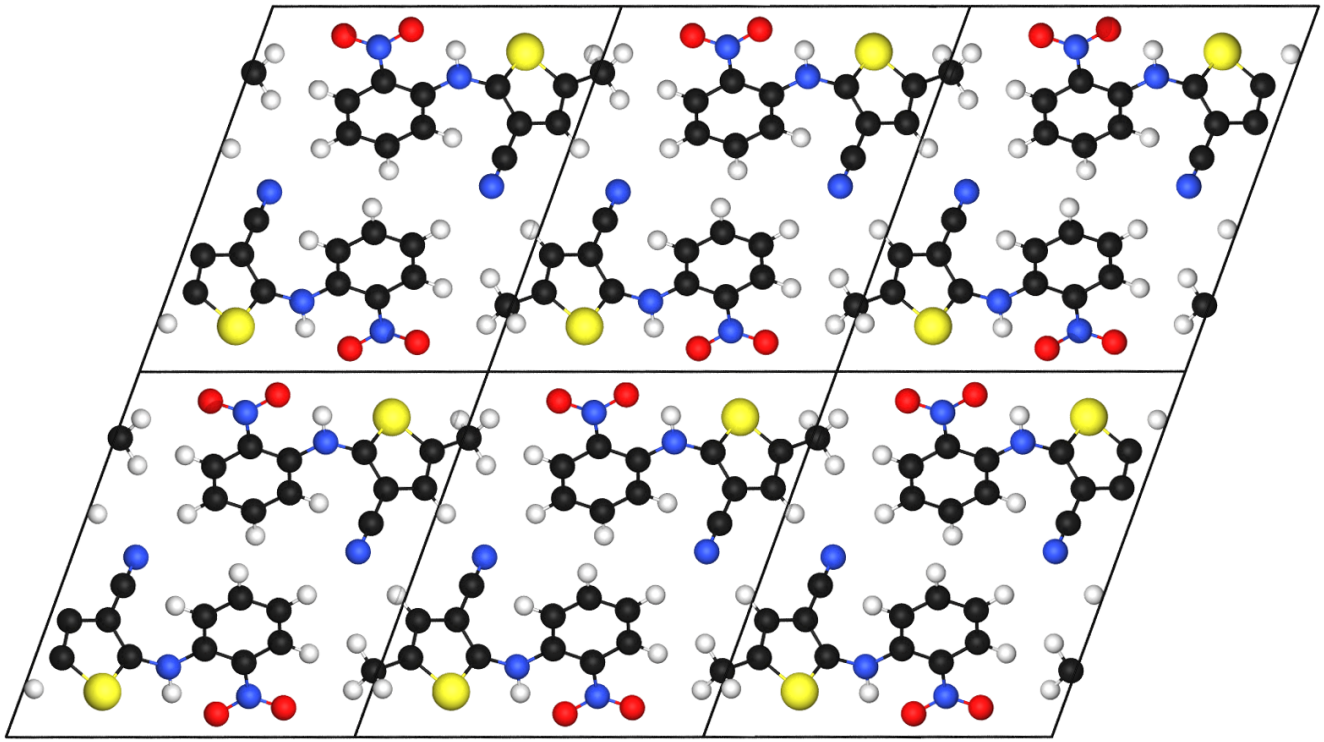
- save computational time since core electrons are unreactive
- include scalar relativistic effects
- ECPs are typically necessary for $Z \geq 37$ (i.e. beyond Kr)
- For transition metals, the outermost s and p electrons should be modeled explicitly, not included in the ECP





Periodic boundary conditions

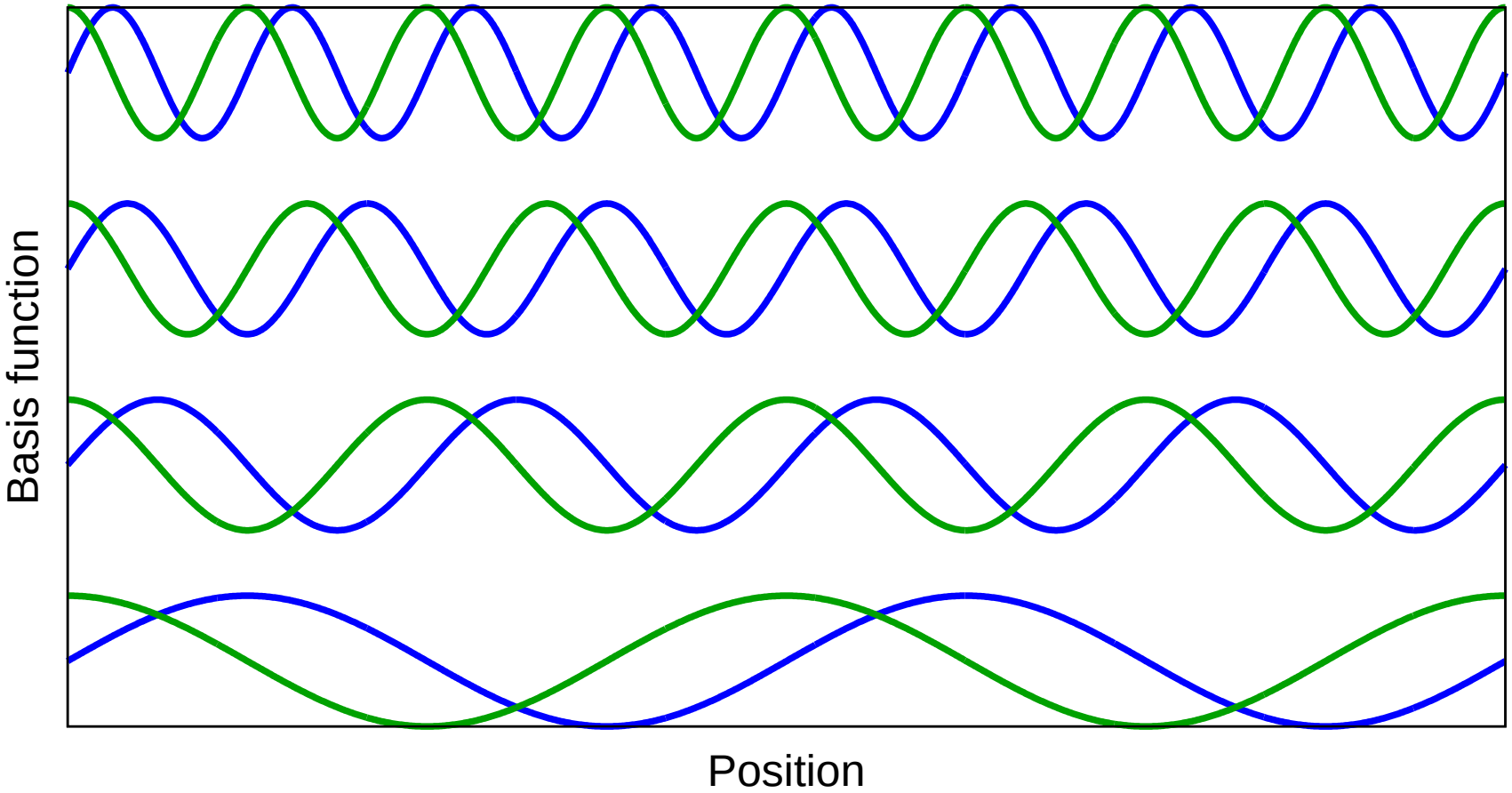
Most solids are crystals than can be represented by a single unit cell, repeatedly replicated in 3D.



The basis functions should have the same periodicity as the lattice.

Plane-wave basis sets

All plane waves are included up to some energy (frequency) cutoff.



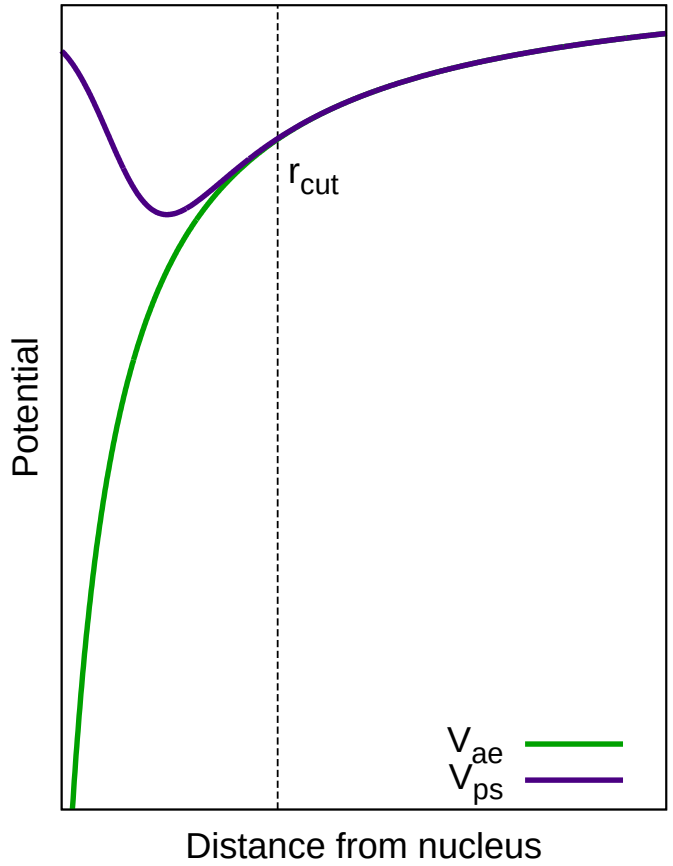
Pseudopotentials

Sharp electron-density peaks near nuclei are hard to represent using plane waves.

Replace the potential (and the density) within some cut-off radius, r_{cut}.

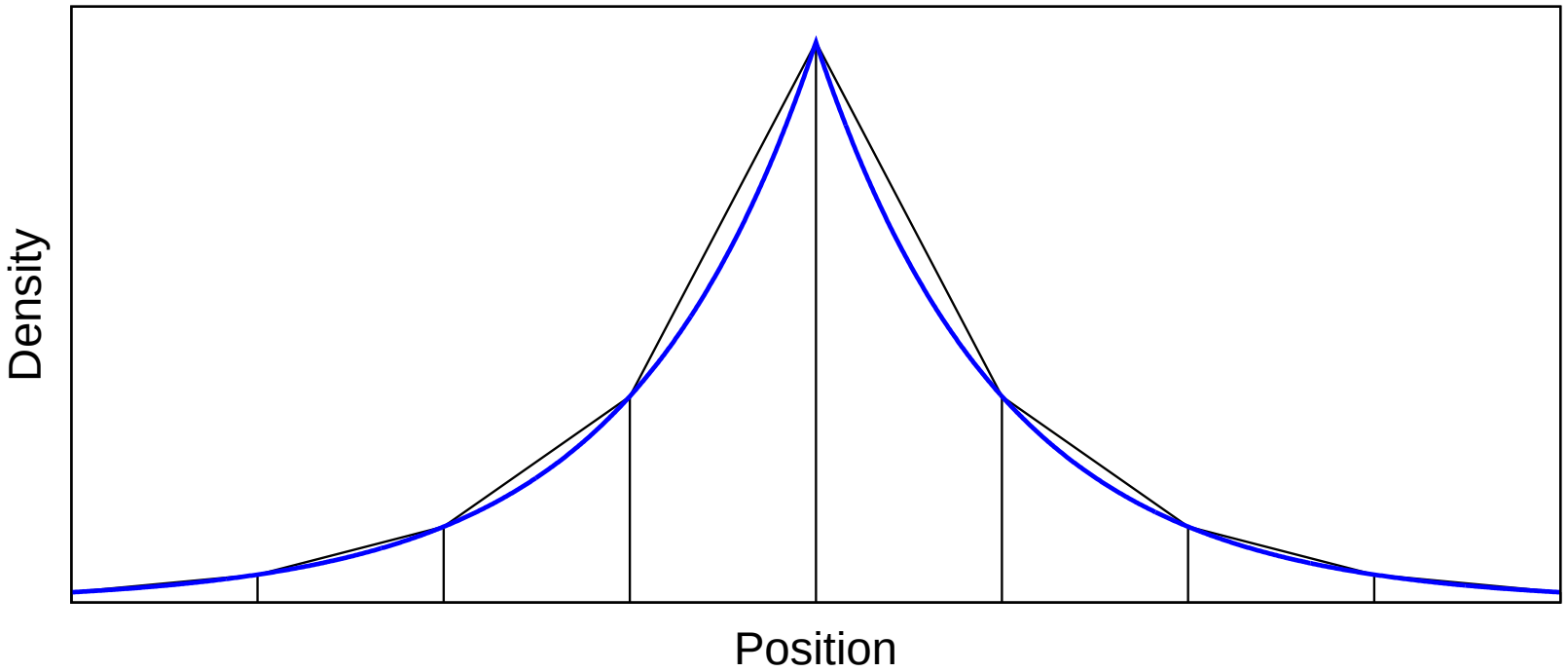
Pseudopotentials are used for all elements.

Core regions between bonded atoms must never overlap.



Integrals: k-point mesh

Integrals to evaluate the energy are done numerically – need a dense enough mesh to obtain converged values.



In reciprocal space – more points for small cell dimensions and few points for large cell dimensions.

Summary

DFT is a powerful tool for chemistry, physics, materials science, and engineering, but is not a black box – **always seek input from experts**.

It is important that enough detail is given for the calculations to be reproducible and for the reader to judge their quality.

- Geometry details for the system under study
  - cif files or xyz coordinates
  - charge and magnetization
- The choice of density functional
  - functional name, citation, dispersion correction
- Basis set or plane-wave/pseudopotential details
  - basis set and ECP
  - pseudopotential type, plane-wave cutoffs, k-point mesh