

Derivatives

change in geometry (R), external electric field (F), external magnetic field (M), nuclear magnetic field (I)

n_R	n_F	n_B	n_I	Property
1	0	0	0	Energy gradient \mathbf{g}
2	0	0	0	Harmonic vibrational frequencies ω_i
0	1	0	0	Electric dipole moment μ
0	2	0	0	Electric polarizability α
0	0	1	0	Magnetic (dipole) moment μ
0	0	2	0	magnetic susceptibility χ
0	0	0	1	ESR hyperfine coupling constant a_i
1	1	0	0	Intensities of fundamental IR transitions
0	1	1	0	Circular dichroism (CD)
0	0	1	1	Nuclear magnetic shielding (\rightarrow chemical shift in NMR)

Gradients

- Molecular forces
 - Commonly implemented
 - Special derivations
- Alchemical derivatives: electronic electrostatic potential

Hessians

- Normal modes
 - Commonly only for spatial derivatives
 - Special derivations (less often)

$f(x)$: total energy $E(R_i, Z_i)$. What if we have no derivatives?

Variants

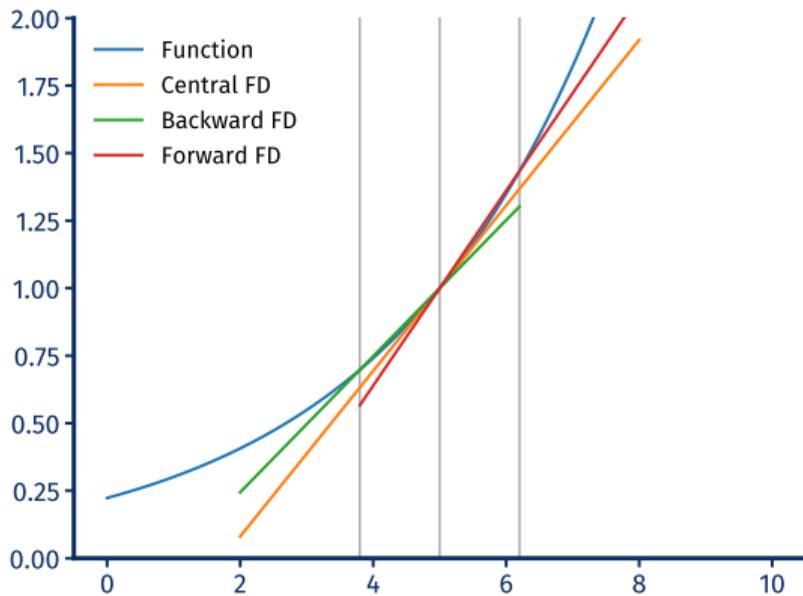
- Forward/backward
- Central
- Higher-order

Issues

- Finite displacement
- Numerical stability / finite precision
- Many calculations

Main advantage

- General applicability
- Just points and weights



Finite differences illustration

Approximative derivatives

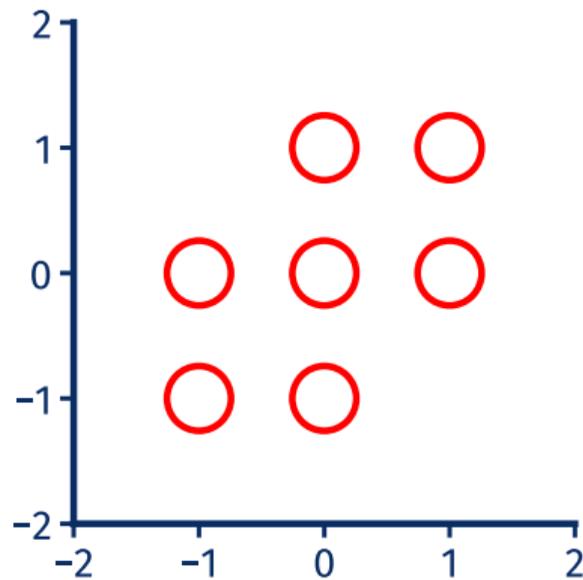
- Similar to the limit expression

Higher dimensions

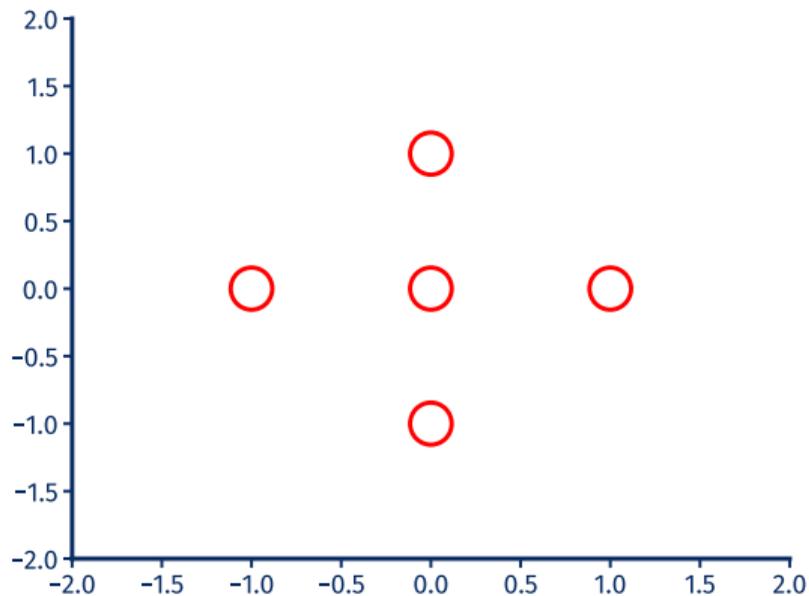
- Even more points
- Set of points: stencil
- Balance quality and cost

Weights

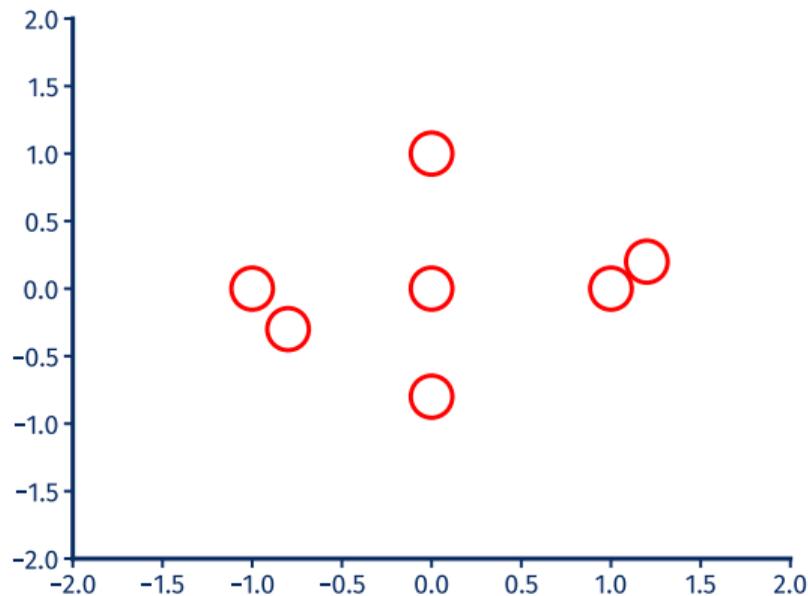
- Choose points first
- Taylor expansion around center
- Solve set of linear equations



2D stencil



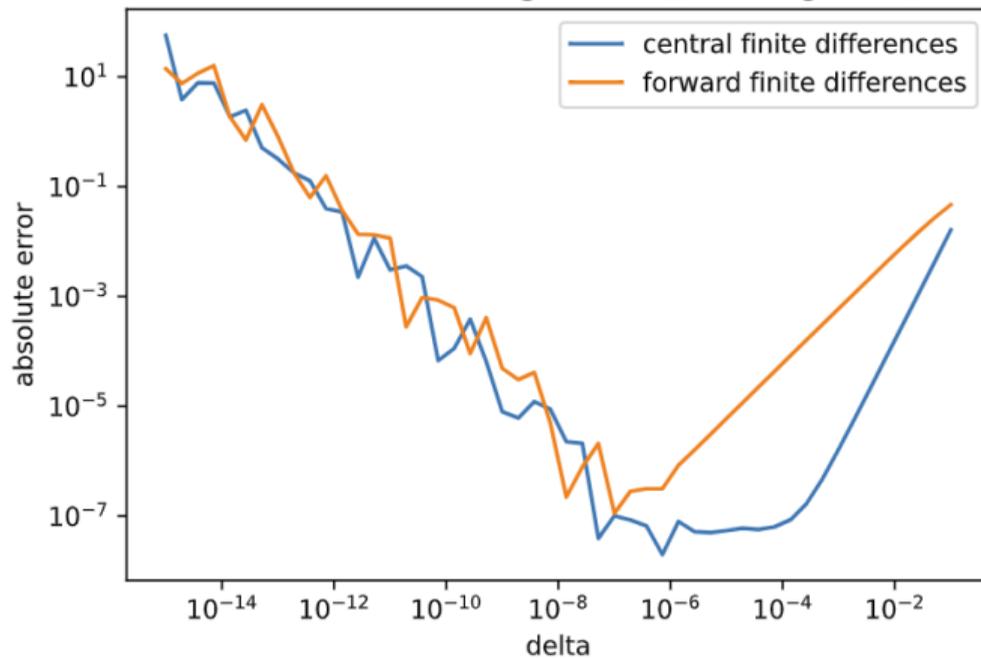
2D stencil



Diagonal stencil

Are both valid for all mixed derivatives?

H-F finite differences vs reference gradient @ 1.0 Angstrom, RHF/cc-pVDZ



$$\frac{dE}{d\mathbf{R}} = \frac{d}{d\mathbf{R}} \langle \psi | \hat{H} | \psi \rangle \quad (1)$$

$$= \left\langle \frac{d\psi}{d\mathbf{R}} \left| \hat{H} \right| \psi \right\rangle + \left\langle \psi \left| \hat{H} \right| \frac{d\psi}{d\mathbf{R}} \right\rangle + \left\langle \psi \left| \frac{d\hat{H}}{d\mathbf{R}} \right| \psi \right\rangle \quad (2)$$

$$= E \left\langle \frac{d\psi}{d\mathbf{R}} \left| \psi \right\rangle + E \left\langle \psi \left| \frac{d\psi}{d\mathbf{R}} \right\rangle + \left\langle \psi \left| \frac{d\hat{H}}{d\mathbf{R}} \right| \psi \right\rangle \quad (3)$$

First two terms cancel if

- Complete basis set
- At least: basis set does not change along derivative

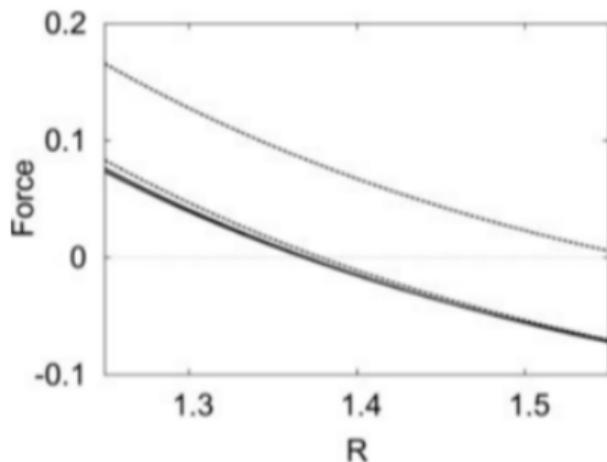


Figure 1. Dashed lines: Hellmann-Feynman force (upper) and energy gradient (lower) in H_2 for a minimal basis set (a single $1s(\zeta = 1.2)$ function per center). Solid lines: same for the basis set augmented with a $2p(\zeta' = 2.0)$ function on each center.

- Complete basis
- Basis does not change along perturbation
 - Includes spatial changes
- Pulay forces / stress

- Hartree-Fock / DFT alone yields energy, but no derivative
- Starting from second order, coupled-perturbed equations need to be solved
 - Response of orbital solutions due to perturbation
- Part of almost any quantum chemistry code
- Wigner's rule: $2n+1$ theorem
 - If you have the wavefunction up to n -th order, you know the perturbed properties up to $2n+1$ th order

- A pure function in programming can be seen as mathematical function
- Programs are nested/chained function calls
- Functions are chained statement
- Idea: use chain rule throughout the code: as long as statements are differentiable
- Limitations:
 - Cannot be done with arbitrary code
 - Crossing boundaries between languages is hard
 - Substantial memory or compute requirements
- Main advantage:
 - No manual derivation needed

- Self-consistent cycle can be a pure function
- Build call tree, apply chain rule
- Technical challenges:
 - Requires re-implementation
 - Crossing language borders is hard
 - Some syntactic sugar is unsupported
 - Slow, can be boosted with "Harris functional" like step
- Chemistry challenges:
 - Some derivatives not well defined (e.g. w.r.t. basis set)

$$z = f(x_1, x_2)$$

$$= x_1 x_2 + \sin x_1$$

$$= w_1 w_2 + \sin w_1$$

$$= w_3 + w_4$$

$$= w_5$$

(4)

(5)

(6)

(7)

(8)

– Most natural

– Compute intensive: needs seed (0, 1) for next variable, so each gradient costs the same as the original function

Operations to compute value	Operations to compute derivative
$w_1 = x_1$	$\dot{w}_1 = 1$ (seed)
$w_2 = x_2$	$\dot{w}_2 = 0$ (seed)
$w_3 = w_1 \cdot w_2$	$\dot{w}_3 = w_2 \cdot \dot{w}_1 + w_1 \cdot \dot{w}_2$
$w_4 = \sin w_1$	$\dot{w}_4 = \cos w_1 \cdot \dot{w}_1$
$w_5 = w_3 + w_4$	$\dot{w}_5 = \dot{w}_3 + \dot{w}_4$

$$z = f(x_1, x_2)$$

$$= x_1 x_2 + \sin x_1$$

$$= w_1 w_2 + \sin w_1$$

$$= w_3 + w_4$$

$$= w_5$$

- More memory, faster compute
- Multiple derivatives at once
- Can be done on
 - other primitives
 - control structures
- Works with matrices

$$(9) \quad \frac{dz}{dz} = 1 \quad (14)$$

$$(10) \quad \frac{dz}{dz} = 1$$

$$(11) \quad z = w_5 \Rightarrow \frac{dz}{dw_5} = 1 \quad (15)$$

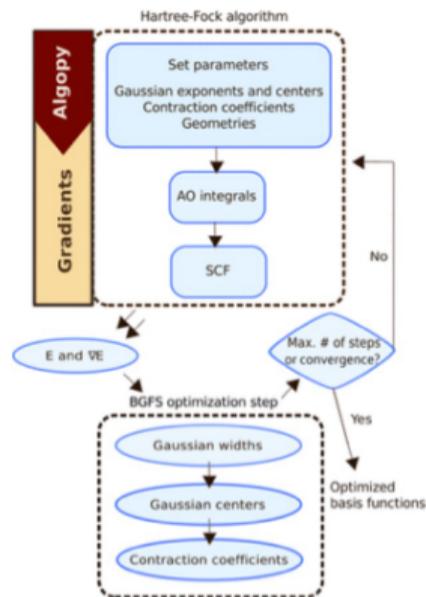
$$(12) \quad \frac{dz}{dw_4} = \frac{dz}{dw_5} \frac{dw_5}{dw_4} = 1 \times 1 = 1 \quad (16)$$

$$(13) \quad \frac{dz}{dw_3} = \frac{dz}{dw_5} \frac{dw_5}{dw_3} = 1 \times 1 = 1 \quad (17)$$

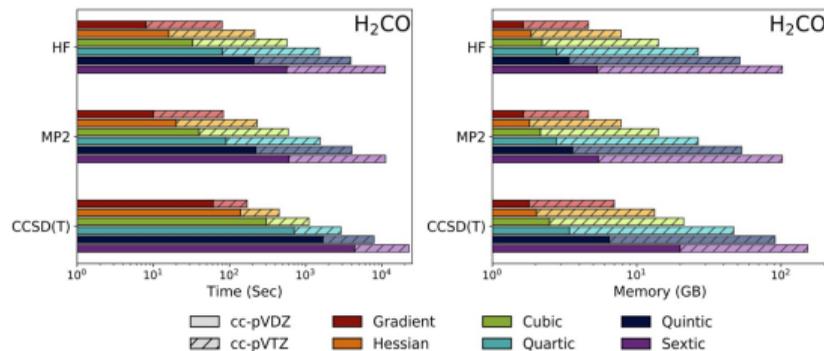
$$\frac{dz}{dw_2} = \frac{dz}{dw_3} \frac{dw_3}{dw_2} = 1 \times w_1 = w_1 \quad (18)$$

$$\frac{dz}{dw_1} = \frac{dz}{dw_3} \frac{dw_3}{dw_1} + \frac{dz}{dw_4} \frac{dw_4}{dw_1} = w_2 + \cos(w_1) \quad (19)$$

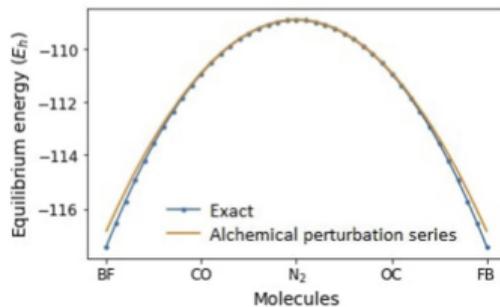
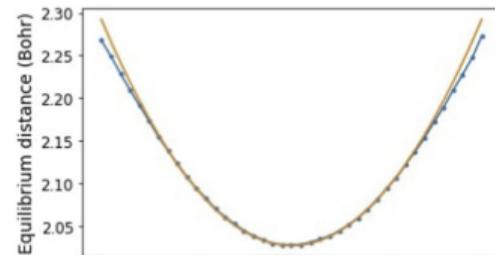
- Forward mode proof-of-concept
- Reason: matrix diagonalisation for degenerate eigenvalues not available in reverse mode
- Optimizes everything
- HF



- Forward mode
- Somewhat stable for degeneracies
- Focuses on spatial derivatives
- HF, MP2, CCSD(T)



- Reverse mode (!)
- Focuses on different properties
- HF, DFT
- Includes alchemical derivatives



Methods for derivatives

- Finite differences: general but costly, needs careful step size selection
- Hellmann-Feynman: exact for complete basis
- Coupled-perturbed: standard in quantum chemistry, analytically derived for each property
- Automatic differentiation: modern approach using chain rule throughout code

Applications

- Energy gradients enable geometry optimization and molecular dynamics
- Higher-order derivatives yield vibrational frequencies, polarizabilities, and spectroscopic properties
- Mixed derivatives provide IR/Raman intensities and response properties
- Alchemical derivatives allow exploration of chemical space