Derivatives and properties

UNIVERSITÄT
DES
SAARLANDES

Stella Stopkowicz

ESQC summer school,

Torre Normanna 2026

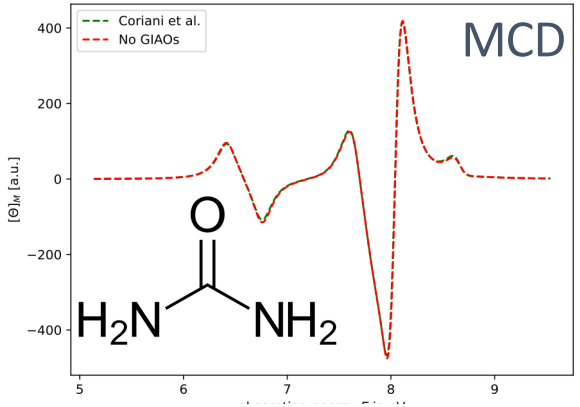
Hylleraas

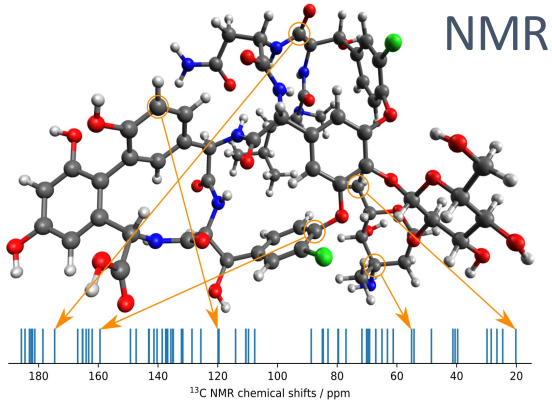
Content

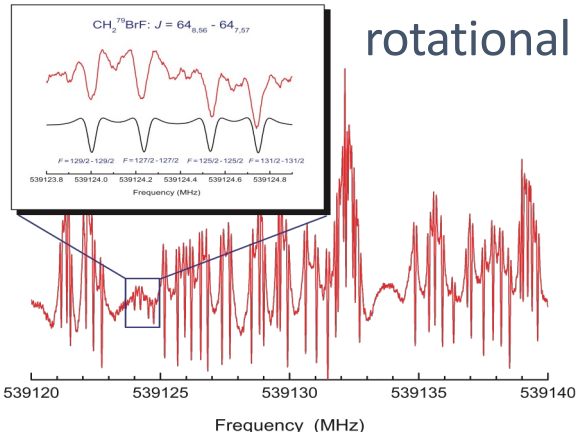
- Part 1: Derivative theory
- Part 2: Magnetic properties and molecules in magnetic fields

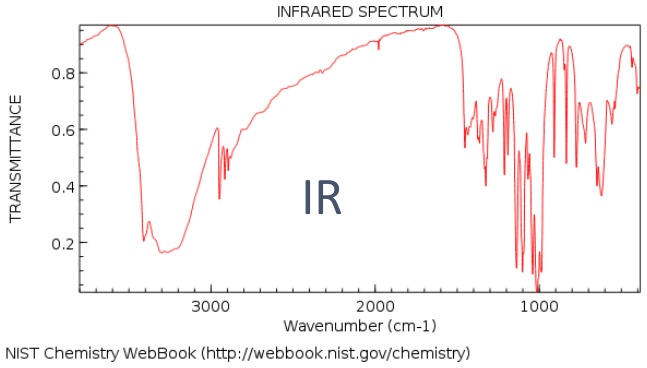
Molecular properties

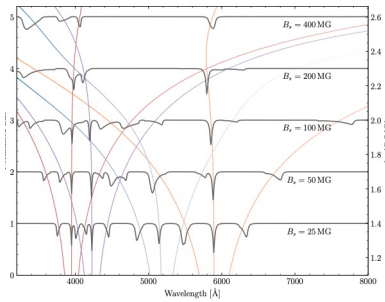
- Spectroscopy
  - NMR, EPR, IR, Raman, Rotational spectra, MCD, …

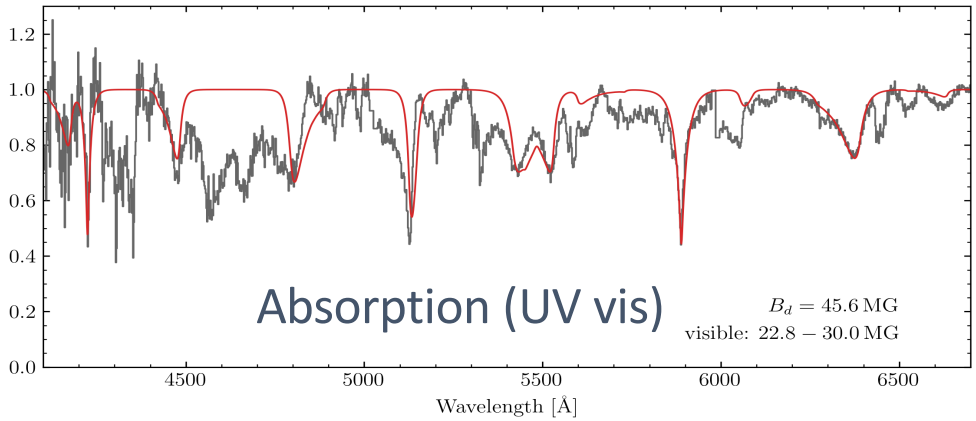












- Time- (or frequency) dependent vs. **time independent**

Types of (time-independent) molecular properties

- State-specific properties
  - Equilibrium structure
  - Vibrational frequencies
  - Dipole moment and polarizability
  - NMR shielding
  - Magnetizability
  - …

Focus of this lecture

- Transition properties
  - Electronic excitation energies
  - One- and two-photon transition strengths
  - Radiative lifetimes
  - Ionization potentials and electron affinities
  - …
- Based on energy differences
  - Reaction energies
  - Stability of conformers/isomers
  - Atomization energies
  - Dissociation energies
  - …

Part 1
Derivative theory

Contents

- Properties as expectation values
- Properties as energy derivatives
  - Numerical and analytical derivatives
- Hellman-Feynman theorem
- Derivative theory
  - Variational and non-variational wave functions
  - Lagrangian formalism
  - (2n+1) and (2n+2) rules
  - Examples for first- and second-order properties via Lagrangian approach

Calculation of molecular properties

Quantum mechanics suggest:

Calculate properties as

expectation value to a corresponding Hermitian operator $\hat{O}$!

$$\frac{\langle \Psi \mid \hat{O} \mid \Psi \rangle}{\langle \Psi \mid \Psi \rangle}$$

Molecular properties as expectation value

- Example: dipole moment
- Calculation via expectation value of respective operator

$$\hat{\boldsymbol{\mu}} = \sum_{\alpha}(-e\mathbf{r}_{\alpha}) + \sum_{A} Z_A \mathbf{R}_A$$

electronic part

nuclear part

$$\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_{\mathrm{el}} + \hat{\boldsymbol{\mu}}_{\mathrm{nuc}}$$

- For quantum-chemical calculation: el. part relevant
- Expectation value: $\langle \Psi \mid \hat{\boldsymbol{\mu}}_{\mathrm{el}} \mid \Psi \rangle$

WF normalized

Molecular properties as energy derivatives

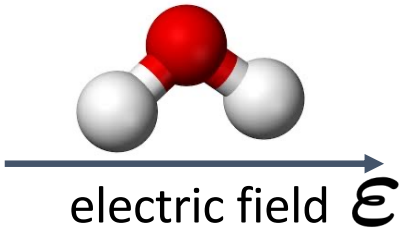
Alternative view: Property as a response of a system to a (small) perturbation

→ to be calculated as energy derivative with respect to a perturbation parameter to which to property is associated

→ sufficiently weak perturbation: Taylor expansion around x=0

Molecule in electric field

- Response of the system to a perturbation



→ $E(\varepsilon)$

Energy of the molecule changes and is a function of the electric field

Example: Perturbation by electric field

- Expansion with respect to electric field as perturbation

$$E(\boldsymbol{\varepsilon}) = E(\boldsymbol{\varepsilon}=0) - \boldsymbol{\mu}^T\boldsymbol{\varepsilon} - \frac{1}{2}\boldsymbol{\varepsilon}^T\boldsymbol{\alpha}\boldsymbol{\varepsilon}\ldots$$

field-free energy

field – dependent terms

expansion coefficients characteristic of the molecule and its quantum states → molecular properties.

- Taylor expansion around the field-free case

$$E(\boldsymbol{\varepsilon}) = E(\boldsymbol{\varepsilon}=0) + \left.\frac{dE}{d\boldsymbol{\varepsilon}}\right|^T_{\varepsilon=\mathbf{0}}\boldsymbol{\varepsilon} + \frac{1}{2}\boldsymbol{\varepsilon}^T\left.\frac{d^2E}{d\boldsymbol{\varepsilon}d\boldsymbol{\varepsilon}}\right|_{\varepsilon=\mathbf{0}}\boldsymbol{\varepsilon}\ldots$$

- Comparison:

$$\boldsymbol{\mu} = -\left.\frac{dE}{d\boldsymbol{\varepsilon}}\right|_{\varepsilon=\mathbf{0}}$$

dipole moment

$$\boldsymbol{\alpha} = -\left.\frac{d^2E}{d\boldsymbol{\varepsilon}d\boldsymbol{\varepsilon}}\right|_{\varepsilon=\mathbf{0}}$$

polarizability

Expectation value vs. derivative

- We have now seen two ways to calculate the dipole moment:

expectation value

$$\boldsymbol{\mu} = \langle \Psi \mid \hat{\boldsymbol{\mu}} \mid \Psi \rangle$$

vs.

energy derivative

$$\boldsymbol{\mu} = - \left. \frac{dE}{d\boldsymbol{\varepsilon}} \right|_{\varepsilon = \mathbf{0}}$$

Expectation value vs. derivative

- Equivalence? → Hellmann-Feynman theorem!

$$\frac{dE}{dx} = \langle \Psi \mid \frac{d\hat{H}}{dx} \mid \Psi \rangle$$

WF normalized:
$\langle \Psi \mid \Psi \rangle = 1$

with $\hat{H} = \hat{H}_0 - \hat{\boldsymbol{\mu}}\boldsymbol{\varepsilon}$

$\hat{\boldsymbol{\mu}}\boldsymbol{\varepsilon}$ modifies V_{ne} such that

$$-\frac{dE}{d\boldsymbol{\varepsilon}} = \langle \Psi \mid \hat{\boldsymbol{\mu}} \mid \Psi \rangle = \boldsymbol{\mu}$$

→ Dipole moment as energy derivative

→ Equivalence to perturbation theory expression

Proof of Hellmann-Feynman theorem for exact wave function

- Derivative of energy-expectation value:

only exact WF that solves Schrödinger equation

$$\frac{dE}{dx} = \frac{d}{dx}\langle\Psi \mid \hat{H} \mid \Psi\rangle = \langle\Psi \mid \frac{d\hat{H}}{dx} \mid \Psi\rangle + \underbrace{\left(\langle\frac{d\Psi}{dx} \mid \overbrace{\hat{H}}^{E} \mid \Psi\rangle + c.c.\right)}_{E\left(\langle\frac{d\Psi}{dx} \mid \Psi\rangle + c.c.\right)} \underbrace{E\left(\frac{d}{dx}\langle\Psi \mid \Psi\rangle\right)}_{0}$$

WF normalized

$$\rightarrow \quad \boxed{\frac{dE}{dx} = \langle\Psi \mid \frac{d\hat{H}}{dx} \mid \Psi\rangle}$$

Implications of Hellmann-Feynman theorem

- Only derivatives of Hamiltonian with respect to the perturbation needed (integrals)
- No derivatives of wave function needed

Generally true?

What about approximate wave functions?

Example: Hartree-Fock wave function

Infinitesimal perturbation

$$\Psi_{\mathrm{HF}} \rightarrow \Psi_{\mathrm{HF}} + \delta\Psi_{\mathrm{HF}} = \Psi_{\mathrm{HF}} + \frac{\partial \Psi_{\mathrm{HF}}}{\partial \alpha} \partial\alpha$$

HF variation:

$$E_{\mathrm{HF}} \rightarrow E_{\mathrm{HF}} + \delta E_{\mathrm{HF}}$$

Normalized WF

$$= E_{\mathrm{HF}} + \langle \delta\Psi_{\mathrm{HF}} \mid \hat{H} \mid \Psi_{\mathrm{HF}} \rangle + \langle \Psi_{\mathrm{HF}} \mid \hat{H} \mid \delta\Psi_{\mathrm{HF}} \rangle$$

$$\underbrace{\left(\langle \frac{\partial \Psi_{\mathrm{HF}}}{\partial \alpha} \mid \hat{H} \mid \Psi_{\mathrm{HF}} \rangle + \langle \Psi_{\mathrm{HF}} \mid \hat{H} \mid \frac{\partial \Psi_{\mathrm{HF}}}{\partial \alpha} \rangle \right)}_{=0} \partial\alpha$$

HF variational condition:

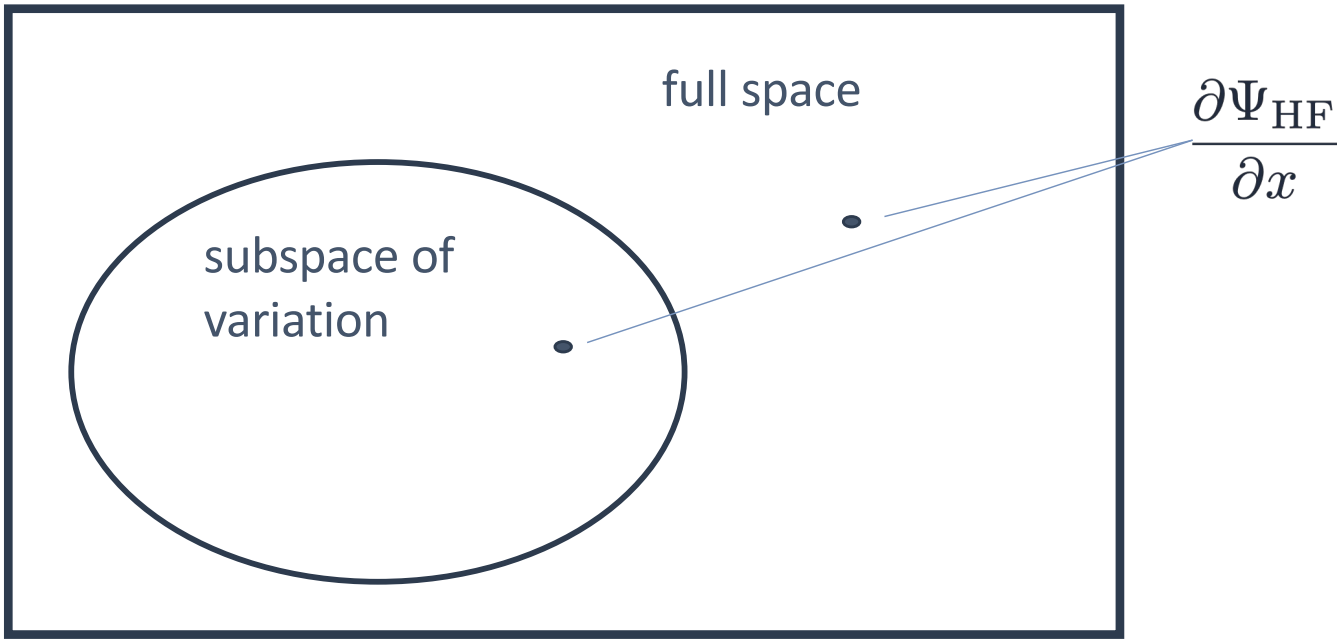
$$\delta E_{\mathrm{HF}} = 0$$

No derivative contributions in HF wave function

→ Hellman-Feynman theorem valid (in the basis-set limit)

Example: Hartree-Fock wave function

- Why Hellman-Feynman theorem not necessarily hold outside the basis-set limit for HF wave function?
- Incomplete basis → variation is being performed within a subspace
- Proof only valid if derivative and variation within the same space



Example: Hartree-Fock wave function

- Hellmann-Feynman theorem valid for Hartree-Fock wave function in the basis-set limit due to variational nature of theory
- Outside the basis-set limit: Does perturbation change metric?

$$\delta_{pq} = \langle \phi_p \mid \phi_q \rangle = \sum_{\mu\nu} c^*_{\mu p} c_{\nu q} S_{\mu\nu}$$

$$S_{\mu\nu} = \langle \chi_\mu \mid \chi_\nu \rangle$$

$$e^{(\underline{r}-\underline{R}_A)^2}$$

differentiate

$$0 = \sum_{\mu\nu} \left(\frac{\partial (c^*_{\mu p} c_{\nu q})}{\partial x} S_{\mu\nu} + c^*_{\mu p} c_{\nu q} \frac{\partial S_{\mu\nu}}{\partial x} \right)$$

Hellmann-Feynman theorem valid if $\frac{\partial S_{\mu\nu}}{\partial x} = 0$

→ If basis functions do not depend on perturbation → no change of the metric → Hellmann-Feynman theorem valid

✅ electric field (for dipole moment)

❌ nuclear coordinates (for geometric perturbations)

Derivative theory

→ Useful concept: we can calculate properties as energy derivatives

Examples of derivatives and their connection to properties

Response to

- Geometrical perturbations
  Forces and force constants
- External electric fields
  Permanent and induced moments, vibrational intensities
- Nuclear quadrupole moments
  Nuclear field gradients, quadrupole-coupling constants
- Magnetic fields
  magnetizabilities
- Magnetic fields and nuclear magnetic moments
  NMR and ESR parameters
- Magnetic and electric fields
  Magnetic circular dichroism, optical rotation
- Molecular rotation
  Spin-rotation constants and molecular g values

Derivative	Observable
$\frac{dE}{d\varepsilon_i}$	dipole moment; in a similar manner also multipole moments, electric field gradients, etc.
$\frac{d^2E}{d\varepsilon_\alpha d\varepsilon_\beta}$	polarizability
$\frac{d^3E}{d\varepsilon_\alpha d\varepsilon_\beta d\varepsilon_\beta}$	(first) hyperpolarizability
$\frac{dE}{dx_i}$	forces on nuclei; stationary points on potential energy surfaces, equilibrium and transition state structures
$\frac{d^2E}{dx_i dx_j}$	harmonic force constants; harmonic vibrational frequencies
$\frac{d^3E}{dx_i dx_j dx_k}$	cubic force constants; vibrational corrections to distances and rotational constants
$\frac{d^4E}{dx_i dx_j dx_k dx_l}$	quartic force constants; anharmonic corrections to vibrational frequencies
$\frac{d^2E}{dx_i d\varepsilon_\alpha}$	dipole derivatives; infrared intensities within the harmonic approximation
$\frac{d^3E}{dx_i d\varepsilon_\alpha d\varepsilon_\beta}$	polarizability derivative; Raman intensities
$\frac{d^2E}{dB_\alpha dB_\beta}$	magnetazibility
$\frac{d^2E}{dm_{K\,j} dB_\alpha}$	nuclear magnetic shielding tensor; relative NMR shifts
$\frac{d^2E}{dI_{K\,i} dI_{L\,j}}$	indirect spin-spin coupling constant
$\frac{d^2E}{dB_\alpha dJ_\beta}$	rotational g-tensor; rotational spectra in magnetic field
$\frac{d^2E}{dI_{K\,i} dB_\alpha}$	nuclear spin-rotation tensor; fine structure in rotational spectra
$\frac{dE}{dm_{K\,j}}$	spin density; hyperfine interaction constants
$\frac{d^2E}{dS_i dB_\alpha}$	electronic g-tensor

and many more

J. Gauss, Modern Methods and Algorithms in Quantum Chemistry, 3, 541, (2000)

Potential-energy surfaces & geometrical derivatives

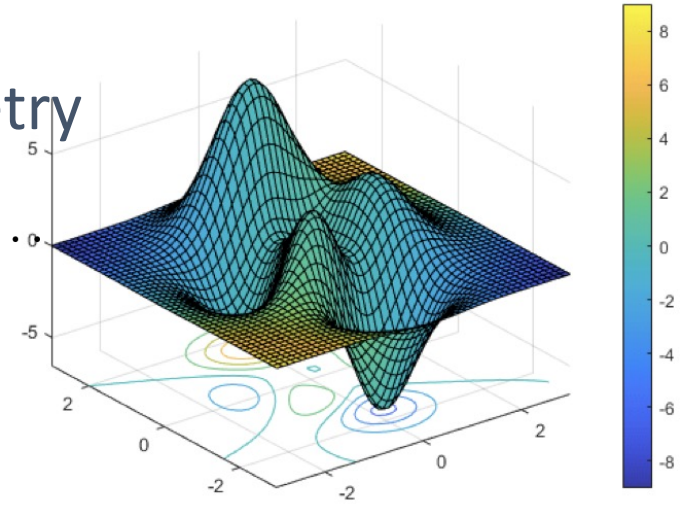
- In Born-Oppenheimer approximation, the potential-energy surface (PES) is a function of the molecular geometry

- Taylor expansion around a reference geometry

$$E(x) = E(x_0) + \left.\frac{dE}{dx}\right|_{x_0} (x - x_0) + \frac{1}{2} \left.\frac{d^2E}{dx^2}\right|_{x_0} (x - x_0)^2 + \ldots$$

molecular gradient → forces

molecular hessian → vib. frequencies



- Needed to locate and characterize critical points
  - Minima → (meta-)stable geometries
  - Saddle points → transition states

- Relation to IR → vibrational frequencies, spectroscopic constants, intensities

Fig: https://www.mathworks.com/help/matlab/ref/surfc.html

NMR parameters

Expansion of energy with respect to magnetic moments *m* and magnetic field *B*

(due to interaction with vector potential, see also part 2 on magnetic properties)

$$E(m,B) = E_0 + \frac{dE}{dm}m + \frac{dE}{dB}B + \frac{1}{2}\frac{d^2E}{dBdB}B^2 + \frac{1}{2}\frac{d^2E}{dmdm}m^2 + \frac{d^2E}{dBdm}Bm + \ldots$$

related to magnetizabilites

related to spin-spin coupling

related to NMR shielding (coupling between B and m)

How to calculate the energy derivatives?

In principle two ways:
Numerical vs. analytical derivatives

Numerical differentiation

- Numerical differentiation (symmetric form)

$$\left.\frac{dE}{dx}\right|_{x_0} \approx \frac{E(x_0+\Delta x)-E(x_0-\Delta x)}{2\Delta x}$$

looks rather simple…, “no” new code needed

Why bother doing anything but this?

Important questions:

- What does it cost?
- How accurate is it?

Cost of numerical differentiation: Gradient

$$\left.\frac{dE}{dx}\right|_{x_0} \approx \frac{E(x_0+\Delta x)-E(x_0-\Delta x)}{2\Delta x}$$

Gradient, geometry optimization

- Example: differentiation wrt nuclear coordinates
  - Number of atoms N_{atom}
  - 3 per direction (x,y,z)
  - 2 per derivative

6 x N_{atom} calculations
(more for better approximations)

→ Scales with number of atoms

→ Will become prohibitive for larger molecules

Cost of analytical differentiation: Gradient

- For Hartree Fock, most time-consuming step is derivative of the two-electron integrals with respect to nuclear coordinates (see also later)
- Number of integrals $\langle \mu\sigma \mid \nu\rho \rangle \sim M^4$
- Number nuclear coordinates $\sim 3\ N_{atom}$

} ~$3N_{atom}$ M^4?

$$\frac{\partial \langle \mu\sigma \mid \nu\rho \rangle}{\partial X_A} = \underbrace{\langle \frac{\partial \mu}{\partial X_A} \sigma \mid \nu\rho \rangle}_{\neq 0 \text{ Only if BF } \mu \text{ centered @ } A} + \langle \mu \frac{\partial \sigma}{\partial X_A} \mid \nu\rho \rangle + \dots$$

- Number of derivative integrals is 4 BF (centered @A) x 3 coordinates x M^4 = 12 M^4

Important: HF analytic gradient does NOT scale with number of atoms!

Cost comparison for HF gradient

Numerical differentiation

$6 \times N_{atom} \times M^4$

- Scales with number of atoms

Analytical differentiation

$12 \times M^4$ (plus some extra)

Does not scale with number of atoms

Possible to go to larger molecules when analytical derivatives are available.

Accuracy of numerical differentiation

$$\left.\frac{dE}{dx}\right|_{x_0} \approx \frac{E(x_0 + \Delta x) - E(x_0 - \Delta x)}{2\Delta x}$$

- Accuracy is dependent on step-size Δx
- In the limit of $\Delta x \longrightarrow 0$, we would get the exact results
- Hence one would in principle want to choose Δx *as small as possible*
- However, this in practice not a good idea…

Accuracy of numerical differentiation

Example: Energy converged to 10^{-10} E_h

This is already rather optimistic… Anyways: machine precision ~10^{-15}

→ Error in gradient:

According to 2-point formula

$$\left.\frac{dE}{dx}\right|_{x_0} \approx \frac{E(x_0+\Delta x)-E(x_0-\Delta x)}{2\Delta x}$$

$$\frac{2\cdot 10^{-10}}{2\Delta x}$$

Δx	error
10^{-4}	10^{-6}
10^{-6}	10^{-4}
10^{-8}	10^{-2}

- Δx too small: Round-off errors!
- Even more problems with higher derivatives (see next slides)

Accuracy of numerical differentiation

- Δx too large: Contamination from higher derivatives!
- Taylor expansion of the energy:

$$E(x) = E(x_0) + \left.\frac{dE}{dx}\right|_{x_0} \underbrace{(x - x_0)}_{\Delta x} + \frac{1}{2}\left.\frac{d^2E}{dx^2}\right|_{x_0} (x - x_0)^2 + \frac{1}{3!}\left.\frac{d^3E}{dx^3}\right|_{x_0} (x - x_0)^3 + \dots$$

- Needed: $\left.\frac{dE}{dx}\right|_{x_0} \approx \frac{E(x_0 + \Delta x) - E(x_0 - \Delta x)}{2\Delta x}$

For $x = x_0 + \Delta x$

$$E(x_0 + \Delta x) = E(x_0) + \left.\frac{dE}{dx}\right|_{x_0} (\Delta x) + \frac{1}{2}\left.\frac{d^2E}{dx^2}\right|_{x_0} (\Delta x)^2 + \frac{1}{3!}\left.\frac{d^3E}{dx^3}\right|_{x_0} (\Delta x)^3 + \dots$$

For $x = x_0 - \Delta x$

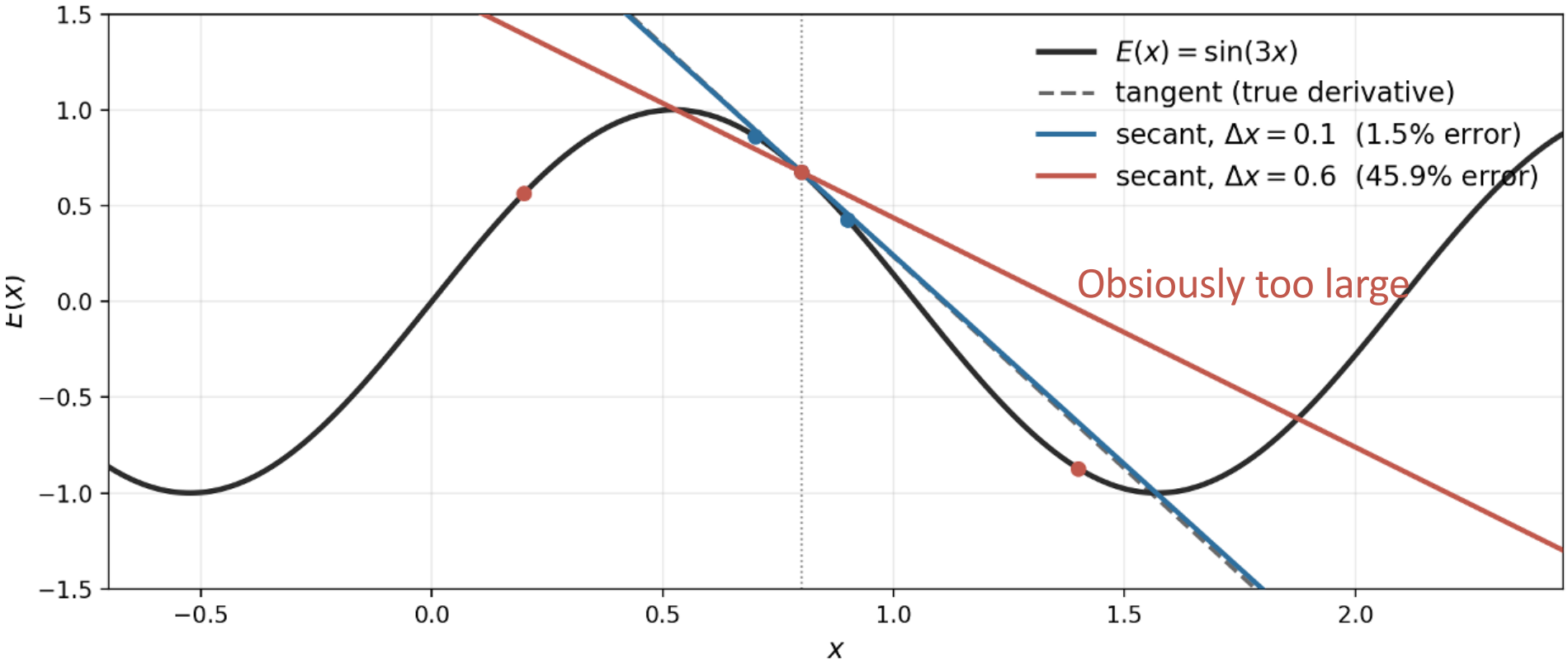
$$E(x_0 - \Delta x) = E(x_0) + \left.\frac{dE}{dx}\right|_{x_0} (-\Delta x) + \frac{1}{2}\left.\frac{d^2E}{dx^2}\right|_{x_0} (-\Delta x)^2 + \frac{1}{3!}\left.\frac{d^3E}{dx^3}\right|_{x_0} (-\Delta x)^3 + \dots$$

→

$$\frac{E(x_0 + \Delta x) - E(x_0 - \Delta x)}{2\Delta x} = \left.\frac{dE}{dx}\right|_{x_0} + \frac{1}{6}\left.\frac{d^3E}{dx^3}\right|_{x_0} \Delta x^2 + \dots$$

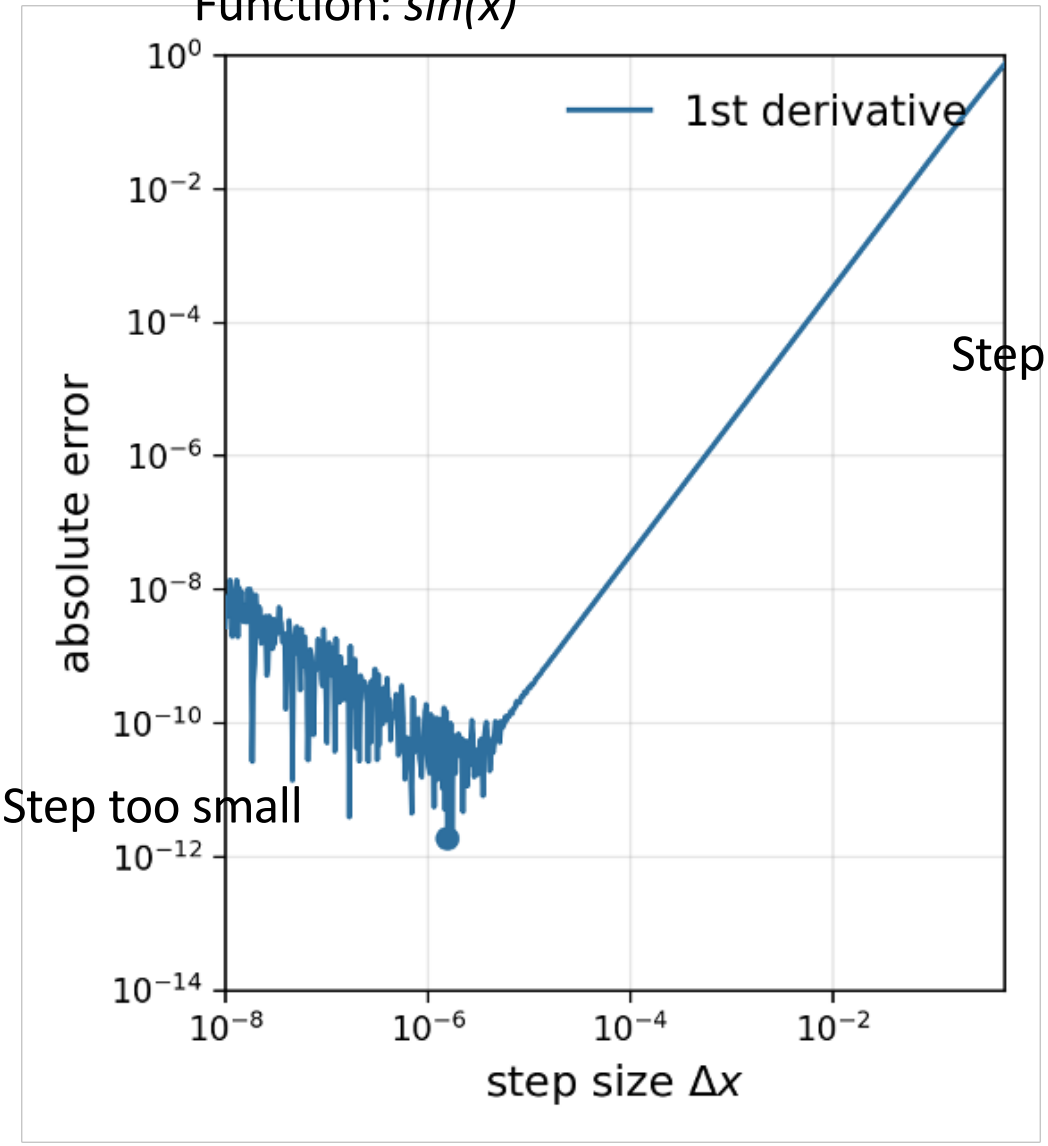
contamination, grows with Δx

Accuracy of numerical differentiation



Accuracy of numerical differentiation

Function: *sin(x)*

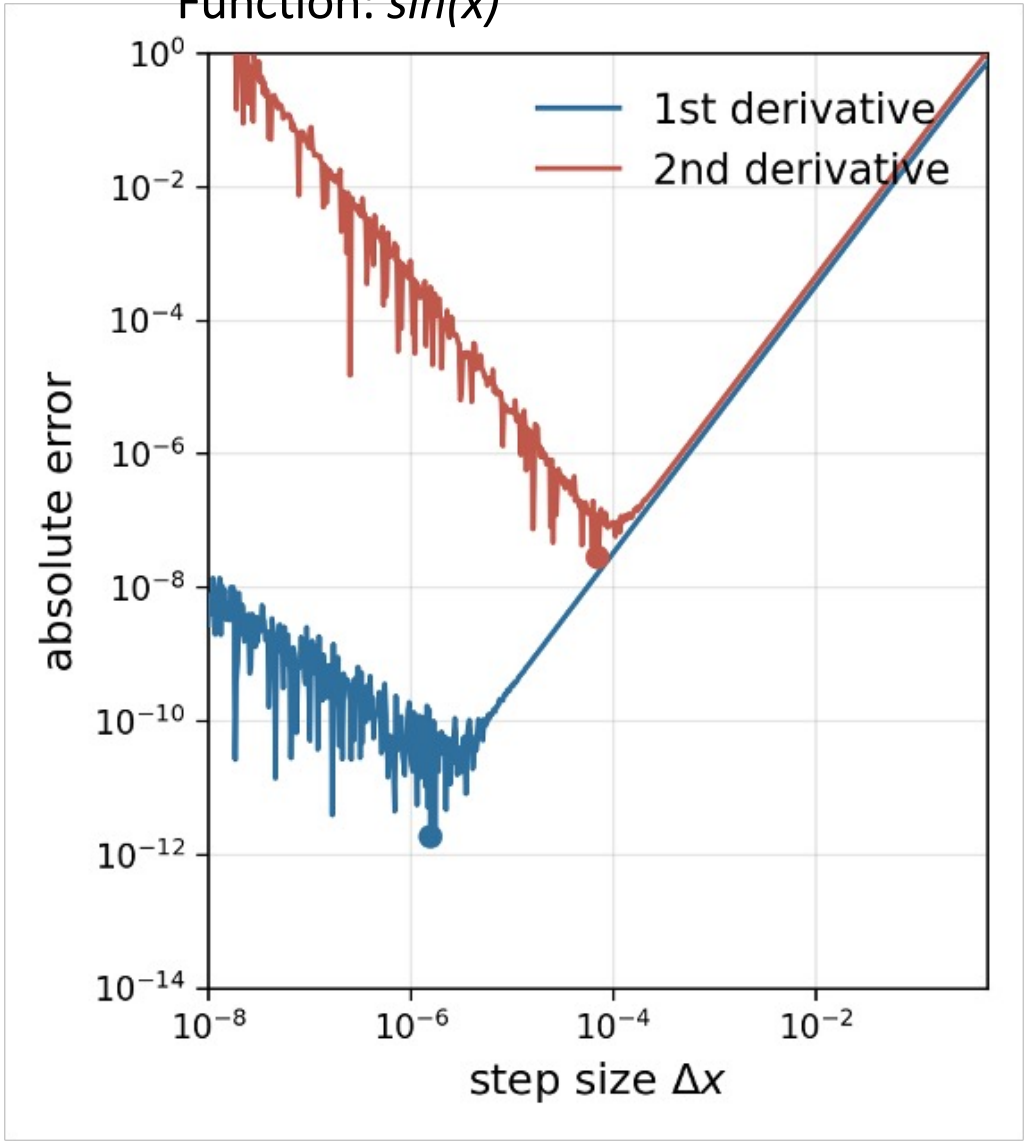


Step too large

Step too small

Accuracy of numerical differentiation

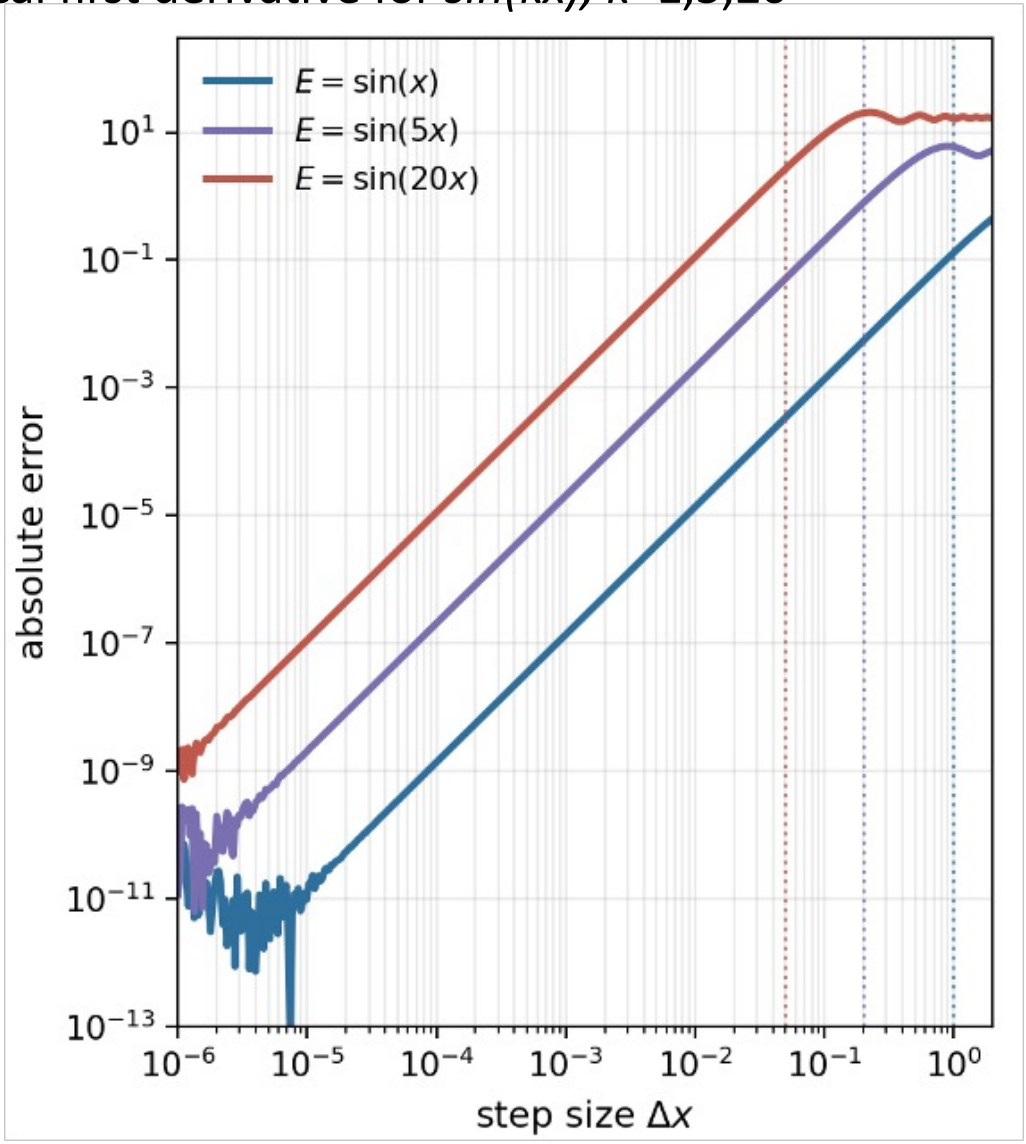
Function: *sin(x)*



Even worse for second derivative!

Accuracy of numerical differentiation

Error in numerical first derivative for *sin(kx), k*=1,5,20



Depends on function!

Accuracy of numerical differentiation

Step size too large: Possible to use higher-order polynomial for the to approximate the derivative but then even more calculations derivative needed:

(*n* points → (*n*-1)th order polynomial)

For example, rather than the 2-point formula

the 4, 6, etc… formula can be used:

2pt: $$b = \frac{f(x) - f(-x)}{2x}$$

4pt: $$b = \frac{8[f(x) - f(-x)] - [f(2x) - f(-2x)]}{12x}$$

6pt: $$b = \frac{45[f(x) - f(-x)] - 9[f(2x) - f(-2x)] + [f(3x) - f(-3x)]}{60x}$$

Accuracy of numerical differentiation

- Or more generally:

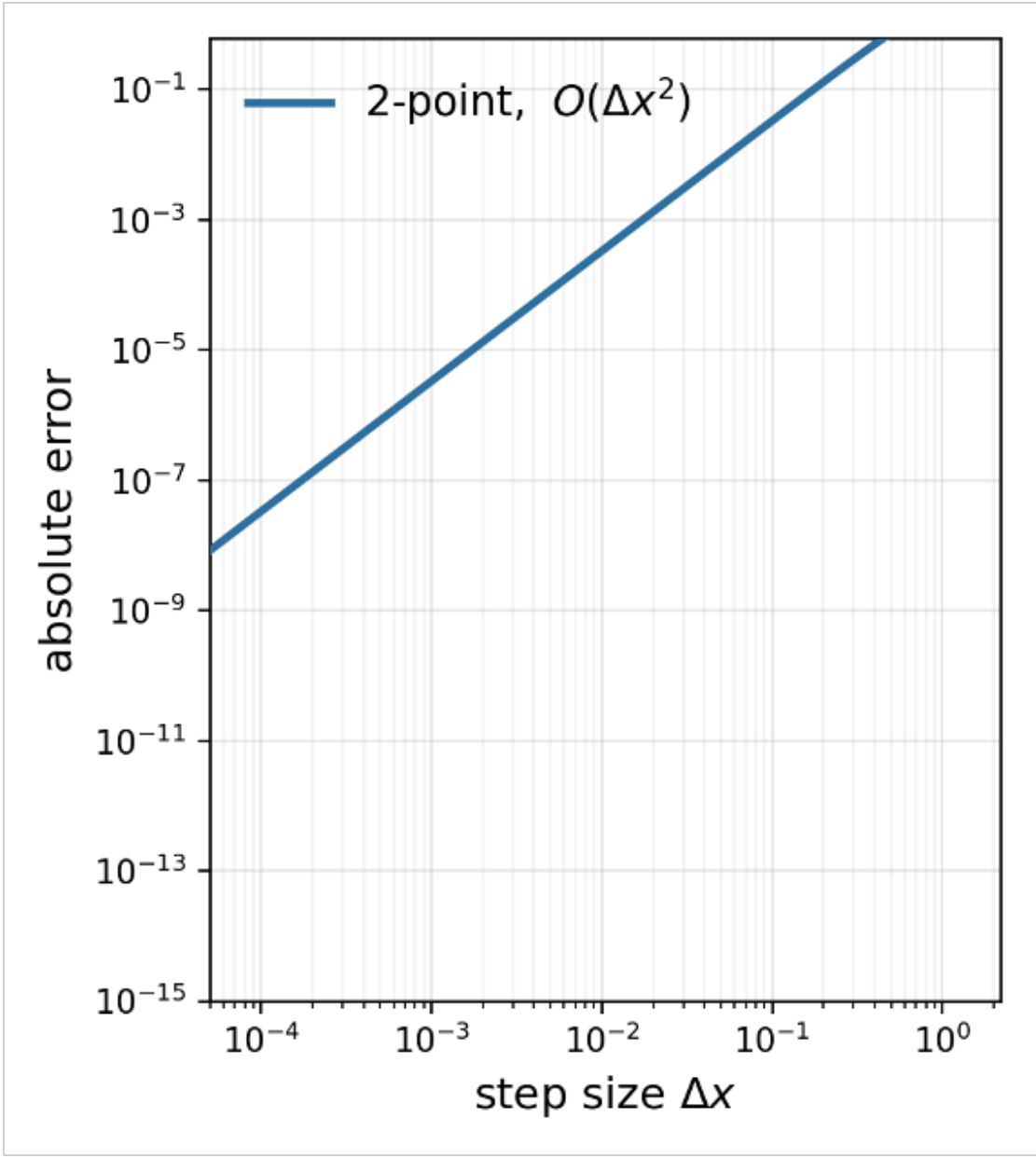
$$f(k) = \underbrace{f(k_0)}_{a} + \underbrace{\left(\frac{\partial f(k)}{\partial k}\right)_{k=k_0}}_{b} \underbrace{(k-k_0)}_{x} + \underbrace{\frac{1}{2}\left(\frac{\partial^2 f(k)}{\partial k^2}\right)_{k=k_0}}_{c} \underbrace{(k-k_0)^2}_{x^2} + \ldots$$

- Solve for coefficients

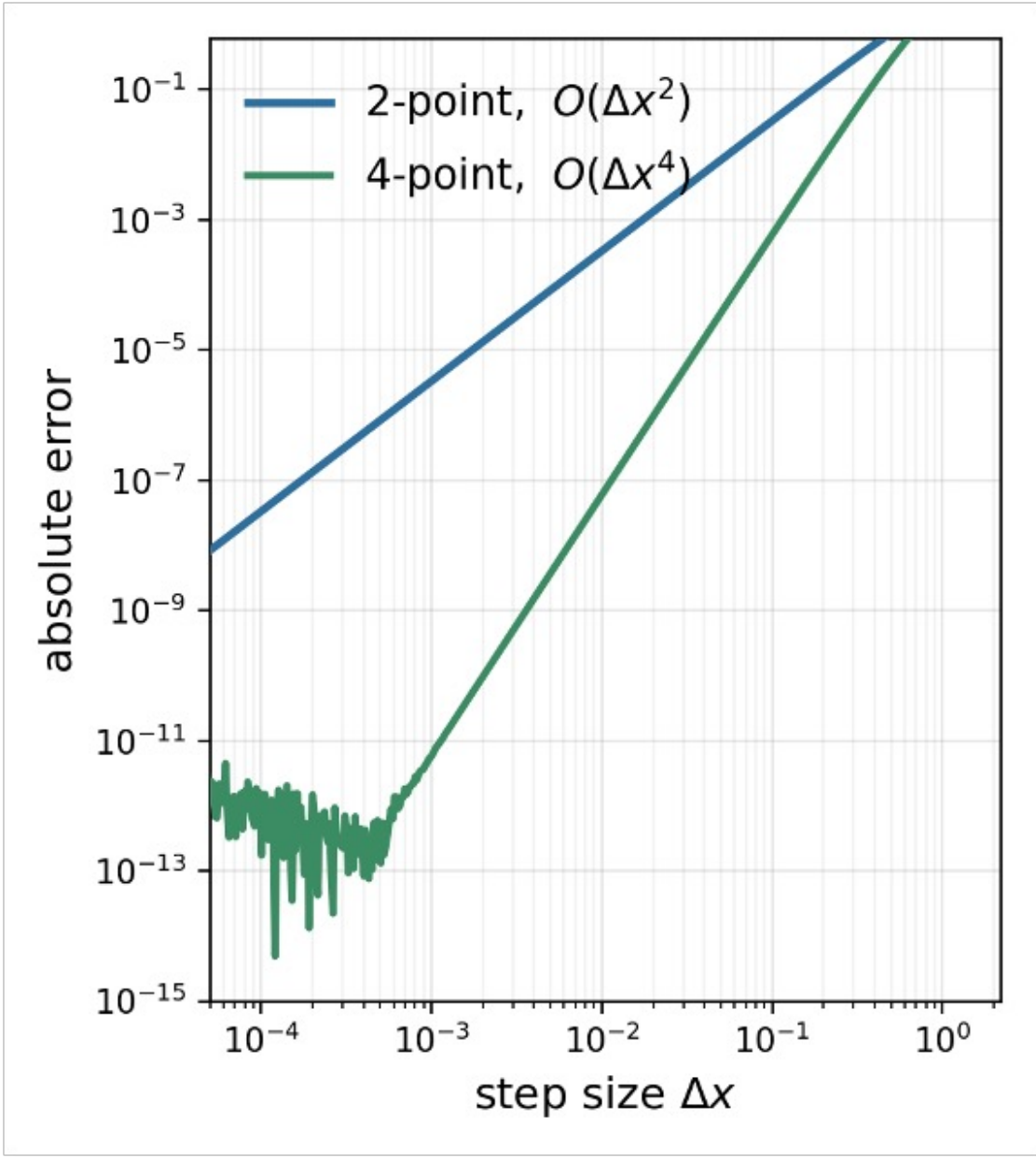
$$\begin{pmatrix} 1 & x & x^2 & x^3 & \ldots \\ 1 & 2x & (2x)^2 & (2x)^3 & \ldots \\ 1 & 3x & (3x)^2 & (3x)^3 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ \vdots \end{pmatrix} = \begin{pmatrix} f(x) \\ f(2x) \\ f(3x) \\ \vdots \end{pmatrix}$$

$$\mathbf{Ac} = \mathbf{f}$$

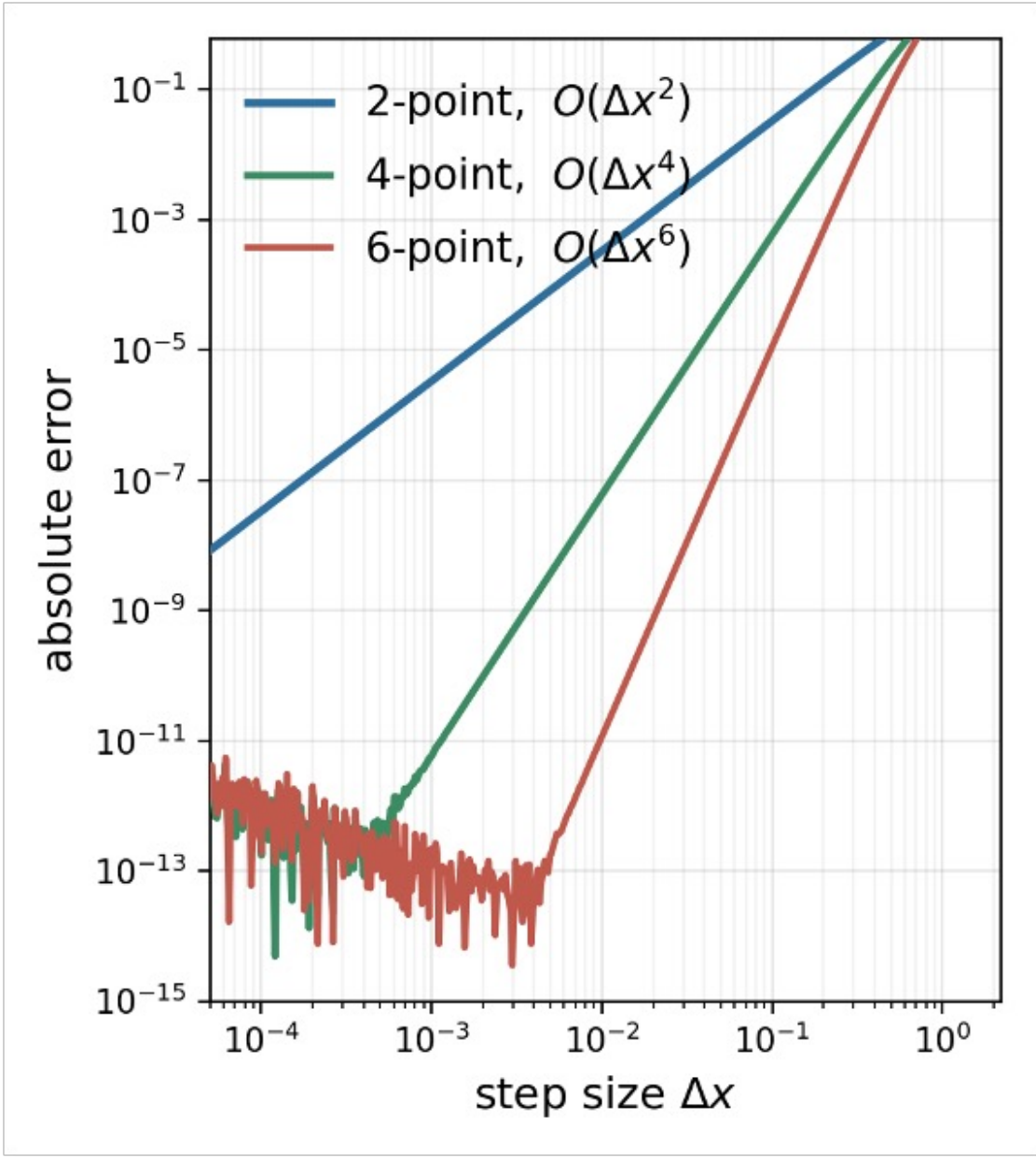
Accuracy of numerical differentiation



Accuracy of numerical differentiation



Accuracy of numerical differentiation



Accuracy of numerical differentiation

Remaining issues:

- More calculations needed for fits to higher-order polynomials
- Range of acceptable step sizes decreases
- Still unclear what a good step size is. This needs calibration (more calculations).

Analytical derivatives

- In principle exact
- Require explicit expressions and their implementation (traditionally new implementation per new method with some re-use)
- Often less computational cost (as compared to numerical differentiation)
- Differences whether or not methods are variational (constrained /unconstrained) or non-variational

Numerical vs analytical differentiation

	numerical	analytical
accuracy	limited	high
efficiency	low	high
freq. dependence	no	yes
imaginary pert.	only with complex wave function	yes
implementation	easy	demanding

→Analytic differentiation preferred if available

Analytical derivatives

Context:

In the following we will deal with analytical derivatives and the „tricks“ to use in order to get the derivatives with as low computational cost as possible

Explicit and implicit dependence on perturbation

- Energy expression, dependence on the perturbation
- Wave function parameters *c* (MO coefficients, amplitudes)

$$E(x) = \langle \Psi(c(x)) \mid \hat{H}(x) \mid \Psi(c(x)) \rangle$$

implicit dependence — explicit dependence — implicit dependence

- In general case need to consider both implicit as well as explicit dependence on perturbation when differentiating, i.e.

$$\frac{dE(x)}{dx} = \frac{\partial E}{\partial x} + \frac{\partial E}{\partial c}\frac{\partial c}{\partial x}$$

$\frac{\partial c}{\partial x}$ → change of WF wrt pertubation x = wave function response

explicit dependence — implicit dependence

Variational wave function

- When WF is determined in unconstrained variation then (as per definition!)

$$\frac{\partial E}{\partial c} \stackrel{!}{=} 0$$

- It follows:

$$\frac{dE(x)}{dx} = \frac{\partial E}{\partial x} + \cancel{\frac{\partial E}{\partial c}\frac{\partial c}{\partial x}}$$

=0

- This means the response of the wave-function parameters $\frac{\partial c}{\partial x}$ does not need to be calculated to obtain the gradient

Examples for variational wave functions

- Hartree-Fock (HF) energy in an unconstrained exponential parametrization

$$| \Psi^{\mathrm{HF}} \rangle = \exp(-\kappa) \mid 0 \rangle, \quad \kappa^{\dagger} = -\kappa$$

- Full CI (FCI) wave function

$$|\Psi^{\mathrm{FCI}}\rangle = \sum_I c_I |\Phi_I\rangle$$

$$E^{\mathrm{FCI}} = \frac{\langle \Psi^{\mathrm{FCI}} \mid \hat{H} \mid \Psi^{\mathrm{FCI}} \rangle}{\langle \Psi^{\mathrm{FCI}} | \Psi^{\mathrm{FCI}} \rangle} \qquad \frac{\partial E^{\mathrm{FCI}}}{\partial c} = 0$$

Constrained variation and non-variational wave functions

- If wave function is constructed via constrained variation or in a non-variational manner, in principle response of wave-function parameters needs to be calculated

$$\frac{dE(x)}{dx} = \frac{\partial E}{\partial x} + \frac{\partial E}{\partial c}\frac{\partial c}{\partial x}$$

doesn't vanish

needs hence to be calculated!

- Possible…., will need a set of equations per perturbation

Example: HF gradient

- HF energy

$$E^{\mathrm{HF}} = \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\sigma\rho} D_{\mu\nu} D_{\sigma\rho} \left(\langle \mu\sigma \mid \nu\rho \rangle - \frac{1}{2} \langle \mu\sigma \mid \rho\nu \rangle \right)$$

- Differentiate

$$\frac{dE(x)}{dx} = \frac{\partial E}{\partial x} + \frac{\partial E}{\partial c} \frac{\partial c}{\partial x}$$

$$\frac{dE^{\mathrm{HF}}}{dx} = \sum_{\mu\nu} D_{\mu\nu} \frac{\partial h_{\mu\nu}}{\partial x} + \frac{1}{2} \sum_{\mu\nu\sigma\rho} D_{\mu\nu} D_{\sigma\rho} \left(\frac{\partial \langle \mu\sigma \mid \nu\rho \rangle}{\partial x} - \frac{1}{2} \frac{\partial \langle \mu\sigma \mid \rho\nu \rangle}{\partial x} \right)$$

$$+ \sum_{\mu\nu} \frac{\partial D_{\mu\nu}}{\partial x} \left\{ h_{\mu\nu} + \sum_{\sigma\rho} D_{\sigma\rho} \left(\langle \mu\sigma \mid \nu\rho \rangle - \frac{1}{2} \langle \mu\sigma \mid \rho\nu \rangle \right) \right\}$$

- would need to solve for perturbed wave-function parameters → **coupled-perturbed HF** (CPHF) equations

Perturbed MO parameters

- CPHF equations from differentiating Brillouin condition

$$\frac{d}{dx} f_{ai} = 0$$

or Roothaan-Hall equations

- Parametrization

$$\frac{\partial c_{\mu r}}{\partial x} = \sum_p c_{\mu p} U^x_{pr}$$

CPHF coefficients

- Linear equations for U^x_{ai}

$$(\varepsilon_a - \varepsilon_i) U^x_{ai} + \sum_b \sum_j (4\langle ab \mid ij\rangle - \langle ab \mid ji\rangle - \langle aj \mid bi\rangle)$$

$$= -f^{(x)}_{ai} + \varepsilon_i S^x_{ai} + \frac{1}{2} \sum_m \sum_j (4\langle am \mid ij\rangle - \langle am \mid ji\rangle - \langle aj \mid mi\rangle) S^x_{mj}$$

$$\underline{\underline{A}}\, \underline{U^x} = \underline{B^x}$$

can be solved iteratively.

cost: N_{pert} x M^5 (integral transformation) → expensive!!!

Perturbed MO parameters

- Do we *actually* need to calculate the perturbed wave-function parameters?

- Solution: enforce stationarity of energy expression with respect to perturbation-dependent parameters *c*

  → method of Lagrange multipliers!

Method of Lagrange multipliers

INTERNATIONAL JOURNAL OF QUANTUM CHEMISTRY, VOL. XXI, 939–940 (1982)

Simple Derivation of the Potential Energy Gradient for an Arbitrary Electronic Wave Function

TRYGVE ULF HELGAKER
Department of Chemistry, University of Oslo, Blindern, Oslo 3, Norway

An analytical expression for the potential energy gradient was first given by Bratož [1] in 1958 for the closed-shell Hartree–Fock case. Over the last few years analytically calculated gradients [2–13] have been extensively used for studying molecular potential energy surfaces, and much credit for this development goes to Pulay [14,15]. The purpose of this letter is to present a very simple derivation of the expression for the gradient for an arbitrary electronic wave function.

Consider an arbitrary wave function $\Psi(p_i)$ which is uniquely determined by the n parameters $\{p_i\}$. The expectation value of the energy E has been minimized with respect to the first m parameters subject to the constraints

$$C_k(p_1,p_2,\ldots,p_n) = 0, \tag{1}$$

i.e.,

$$\frac{\partial E}{\partial p_i} + \sum_k l_k \frac{\partial C_k}{\partial p_i} = 0, \quad i = 1,2,\ldots,m, \tag{2}$$

where l_k are Lagrangian multipliers. Examples of such constraints are the orthonormality of the molecular orbitals for HF functions and the normalization of expansion coefficients for CI functions. We wish to derive an expression for dE/dX, where X is a parameter determining the molecular electronic Hamiltonian H. We note that, since $dC_k/dX = 0$,

$$\frac{dE}{dX} = \frac{dE}{dX} + \sum_k l_k \frac{dC_k}{dX} = \left\langle \frac{dH}{dX} \right\rangle + \sum_{i=1}^{n} \frac{\partial E}{\partial p_i}\frac{\partial p_i}{\partial X} + \sum_k \left(l_k \sum_{i=1}^{n} \frac{\partial C_k}{\partial p_i}\frac{\partial p_i}{\partial X} \right)$$
$$= \left\langle \frac{dH}{dX} \right\rangle + \sum_{i=1}^{n} \left(\frac{\partial E}{\partial p_i} + \sum_k l_k \frac{\partial C_k}{\partial p_i} \right) \frac{\partial p_i}{\partial X}. \tag{3}$$

For each of the optimized parameters the contribution to the sum vanishes due to Eq. (2), whereas for each constant parameter the contribution vanishes due to $\partial p_i/\partial X = 0$. Thus Eq. (3) reduces to a summation over parameters that are neither constant nor optimized. These parameters $\{u_i\}$ may conveniently be termed unstable:

$$\frac{dE}{dX} = \left\langle \frac{dH}{dX} \right\rangle + \sum_i \frac{\partial E}{\partial u_i}\frac{\partial u_i}{\partial X} + \sum_k \left(l_k \sum_i \frac{\partial C_k}{\partial u_i}\frac{\partial u_i}{\partial X} \right). \tag{4}$$

From this expression it is easily seen that all stable functions, in the terminology of Hall [16], obey the Hellmann–Feynman theorem. Conversely, any unstable parameter may give rise to an additional wave function force. Important examples of the latter are orbital coordinates fixed on the atomic nuclei and orbital coefficients in a conventional CI function.

← First paper on this (within quantum-chemical context)

- cited 17 times (according to google scholar)
- Trygve's first single-author paper (pre-PhD)

Better known:

- *A numerically stable procedure for calculating Møller–Plesset energy derivatives, derived using the theory of Lagrangians,* T. Helgaker, P. Jørgensen, and N. C. Handy, *Theor. Chim. Acta* **76**, 227–245 (1989) 124 citations
- *Coupled cluster energy derivatives. Analytic Hessian for the closed-shell coupled cluster singles and doubles wave function: Theory and applications,* H. Koch, H. J. Aa. Jensen, P. Jørgensen, T. Helgaker, G. E. Scuseria, and H. F. Schaefer III, *J. Chem. Phys.* **92**, 4924–4940 (1990) 281 citations

Method of Lagrange multipliers

- Set up energy functional (Lagrangian L) with side condition f
- Side condition written such that

$$f(x, c(x)) = 0$$

- If those side conditions are fulfilled, Lagrangian is equal to the energy itself: $L = E$
- Lagrangian:

$$L = E + \lambda f$$

side condition

Lagrange multiplier

- Enforce stationarity $\frac{\partial L}{\partial c} \stackrel{!}{=} 0$ and $\frac{\partial L}{\partial \lambda} \stackrel{!}{=} 0$

notice: L, not E!

Gives back side condition f

Method of Lagrange multipliers

$$L = E + \lambda f$$

- Stationarity

$$\frac{\partial L}{\partial c} = \frac{\partial E}{\partial c} + \lambda \frac{\partial f}{\partial c} = 0$$

← determines **λ**

Note: perturbation-independent (as compared to $\frac{\partial c}{\partial x}$)

$$\frac{\partial L}{\partial \lambda} = f = 0$$

← determines c

Gives back side condition

- Note: derivative wrt c → **λ**
  derivative wrt **λ** → c

Method of Lagrange multipliers

- Hence for the gradient:

$$\frac{dL(x)}{dx} = \frac{\partial L}{\partial x} + \cancel{\frac{\partial L}{\partial c}\frac{\partial c}{\partial x}} + \cancel{\frac{\partial L}{\partial \lambda}\frac{\partial \lambda}{\partial x}}$$

$\frac{\partial L}{\partial c} = 0$ $\quad$ $\frac{\partial L}{\partial \lambda} = 0$

→We managed to avoid having to solve for response of wave function parameters to perturbation!

$$\frac{dL(x)}{dx} = \frac{\partial L}{\partial x} = \frac{\partial E}{\partial x} + \lambda \frac{\partial f}{\partial x}$$

In addition as compared to expression for variational methods.

Method of Lagrange multipliers

- Constrained variation: side condition corresponds to constraints that need to be fulfilled. Example: orthogonality constraint of the orbitals for Hartree-Fock*

- Non-variational wave functions: side condition corresponds to the conditional equations that are needed to get the wave function parameters (solved anyways)

*note that HF can be formulated in unconstrained manner using orbital rotation

Lagrangian expressions: Hartree Fock

- Hartree-Fock Lagrangian:

$$L^{\mathrm{HF}} = \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\sigma\rho} D_{\mu\nu} D_{\sigma\rho} \left(\langle \mu\sigma \mid \nu\rho \rangle - \frac{1}{2} \langle \mu\sigma \mid \rho\nu \rangle \right) - 2 \sum_{ij} \epsilon_{ij} \left(\sum_{\mu\nu} c^*_{\mu i} S_{\mu\nu} c_{\nu j} - \delta_{ij} \right)$$

E^{HF}

$D_{\mu\nu} = 2 \sum_i c^*_{\mu i} c_{\nu i}$

orthogonality constraint $\langle \phi_i \mid \phi_j \rangle = \delta_{ij}$

Lagrange multipliers (one per constraint)

$$\frac{\partial L^{\mathrm{HF}}}{\partial c} = 0 \quad \rightarrow\rightarrow \text{Hartree-Fock equations}$$

Hartree Fock gradient

- HF gradient:

$$\frac{dE^{\mathrm{HF}}}{dx} = \frac{\partial L^{\mathrm{HF}}}{\partial x} = \sum_{\mu\nu} D_{\mu\nu} \frac{\partial h_{\mu\nu}}{\partial x} + \frac{1}{2} \sum_{\mu\nu\sigma\rho} D_{\mu\nu} D_{\sigma\rho} \left(\frac{\partial \langle \mu\sigma \mid \nu\rho \rangle}{\partial x} - \frac{1}{2} \frac{\partial \langle \mu\sigma \mid \rho\nu \rangle}{\partial x} \right) - \sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x}$$

$$= \frac{\partial E^{\mathrm{HF}}}{\partial x}$$

Energy-weighted density matrix

$$W_{\mu\nu} = 2 \sum_{i} c^{*}_{\mu i} \epsilon_i c_{\nu i}$$

- Essentially only integral derivatives required!
- No derivatives of the density matrices

Coupled Cluster Theory

CC wave function

$$|\Psi_{\mathrm{CC}}\rangle = e^{\hat{T}} \mid \Psi^{\mathrm{HF}}\rangle$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_N$$

$$\hat{T}_1 = \sum_{ai} t_i^a \hat{a}_a^\dagger \hat{a}_i \qquad \hat{T}_2 = \sum_{abij} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i$$

amplitudes t_p

CC equations

Amplitude equations

$$\langle \Psi_p \mid e^{-\hat{T}} \hat{H} e^{\hat{T}} \mid \Psi^{\mathrm{HF}} \rangle = 0$$

CC energy

$$\langle \Psi^{\mathrm{HF}} \mid e^{-\hat{T}} \hat{H} e^{\hat{T}} \mid \Psi^{\mathrm{HF}} \rangle = E^{\mathrm{CC}}$$

Lagrangian expressions: Coupled Cluster

- Coupled-Cluster Lagrangian (non-variational!)

$$L^{\mathrm{CC}} = \langle \Psi^{\mathrm{HF}} \mid \exp(-\hat{T})\hat{H}\exp(\hat{T}) \mid \Psi^{\mathrm{HF}} \rangle + \sum_p \lambda_p \langle \Psi_p \mid \exp(-\hat{T})\hat{H}\exp(\hat{T}) \mid \Psi^{\mathrm{HF}} \rangle$$

E^{CC}

Amplitude equations (determine the CC amplitudes)

Lagrange multipliers (one per constraint)

- Often written as

$$L^{\mathrm{CC}} = \langle \Psi^{\mathrm{HF}} \mid (1+\Lambda)\exp(-\hat{T})\hat{H}\exp(\hat{T}) \mid \Psi^{\mathrm{HF}} \rangle$$

with deexcitation operator $\langle \Psi^{\mathrm{HF}} \mid \Lambda = \sum_p \lambda_p \langle \Psi_p \mid$

Lagrangian expressions: Coupled Cluster

- Full CC Lagrangian and required side conditions (from t-amplitudes and MO-coefficients → orbital relaxation)

$$L^{\mathrm{CC}} = \langle \Psi^{\mathrm{HF}} \mid (1+\Lambda)\exp(-\hat{T})\hat{H}\exp(\hat{T}) \mid \Psi^{\mathrm{HF}} \rangle$$

E + amplitude equations

$$+ \sum_{ai} Z_{ai} f_{ai} + \sum_{pq} I_{pq}(S_{pq} - \delta_{pq})$$

← orbital relaxation part

Brillouin condition (HF minimal condition)

orthogonality condition of MOs

wf parameters 1: CC amplitudes t_p

wf parameters 2: MO coefficients c

→ Total of 3 types of Lagrange multipliers λ_p, Z_{ai}, I_{pq}

Lagrangian expressions: Coupled Cluster

or in terms of density matrices (general form)

$$L^{\mathrm{CC}} = \sum_{pq} \tilde{D}_{pq} f_{pq} + \sum_{pqrs} \Gamma_{pqrs} \langle pq \,||\, rs \rangle + \sum_{pq} I_{pq} S_{pq}$$

Generalized 1-el density matrix
(contains Z_{ai} contribution)

density matrices (contain t_p and λ_p amplitudes)

$$D_{pq} = \langle \Psi^{\mathrm{HF}} \mid (1+\Lambda) e^{-\hat{T}} \{ \hat{a}_p^\dagger \hat{a}_q \} e^{\hat{T}} \mid \Psi^{\mathrm{HF}} \rangle$$

$$\Gamma_{pqrs} = \langle \Psi^{\mathrm{HF}} \mid (1+\Lambda) e^{-\hat{T}} \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \} e^{\hat{T}} \mid \Psi^{\mathrm{HF}} \rangle$$

Lagrangian expressions: Coupled Cluster

- Derivatives with respect to the wave-function parameters

$$\frac{\partial L^{\mathrm{CC}}}{\partial t_p} = 0$$

Λ equations (determine λ_p)

→ Needed to set up (unrelaxed) CC density

$$\frac{\partial L^{\mathrm{CC}}}{\partial c} = 0$$

Z-vector equations (determine Z_{ai}, I_{pq})

→ Perturbation-independent equations

$$\frac{\partial L^{\mathrm{CC}}}{\partial \lambda_p} = 0$$

gives back CC equations

Lagrangian expressions: Coupled Cluster

$$L^{\mathrm{CC}} = \sum_{pq} \tilde{D}_{pq} f_{pq} + \sum_{pqrs} \Gamma_{pqrs} \langle pq \,||\, rs \rangle + \sum_{pq} I_{pq} S_{pq}$$

- Gradient

$$\frac{dE^{\mathrm{CC}}}{dx} = \frac{\partial L^{\mathrm{CC}}}{\partial x} = \sum_{pq} \tilde{D}_{pq} f^{(x)}_{pq} + \sum_{pqrs} \Gamma_{pqrs} \langle pq \,||\, rs \rangle^{x} + \sum_{pq} I_{pq} S^{x}_{pq}$$

Notation: only the integrals (but not MO coefficients are differentiated!

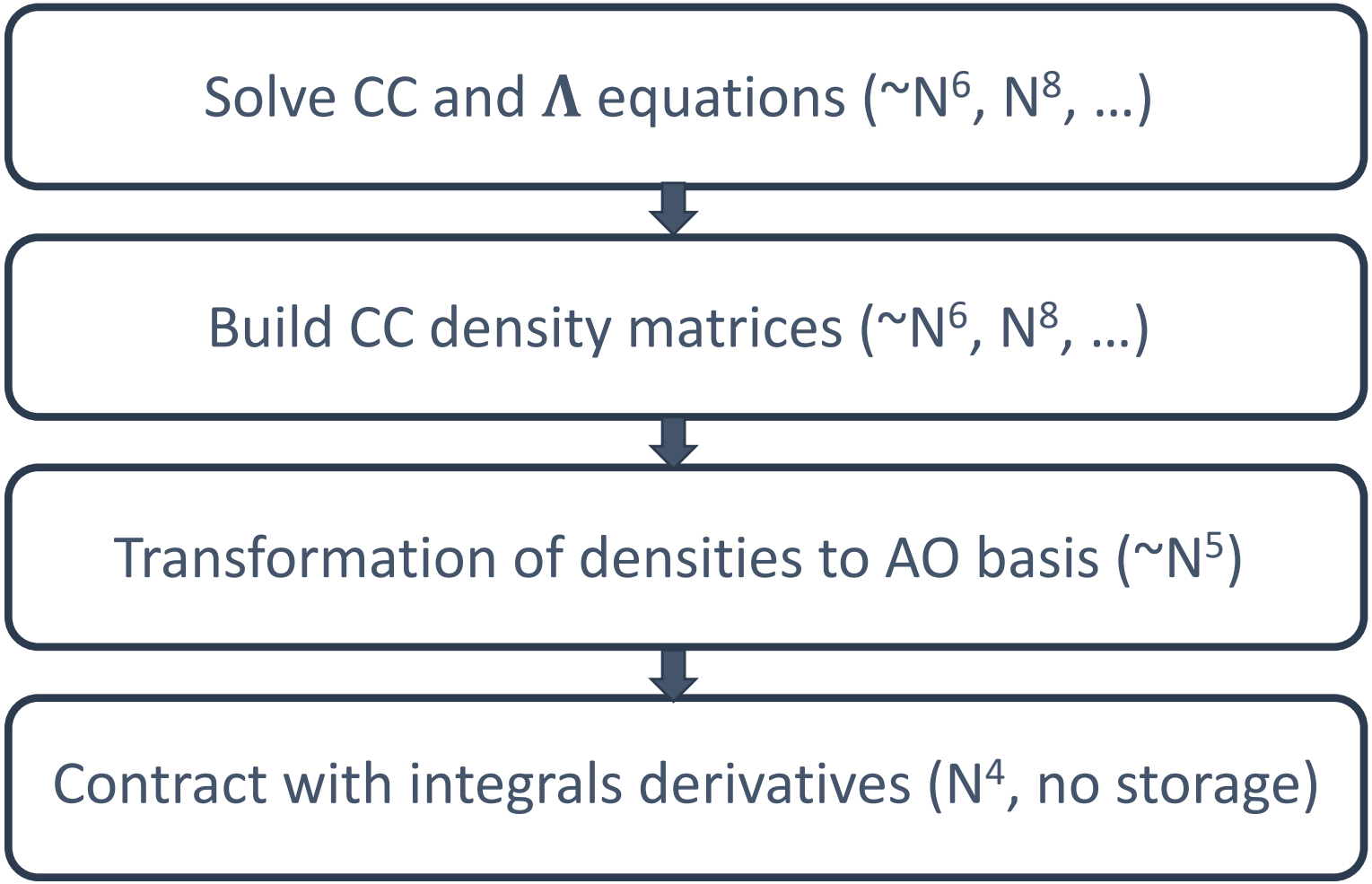
$$f^{(x)}_{pq} = \sum_{\mu\nu} c^{*}_{\mu p} \left\{ \frac{\partial h_{\mu\nu}}{\partial x} + \sum_{\sigma\rho} D^{\mathrm{SCF}}_{\sigma\rho} \frac{\partial \langle \mu\sigma \,||\, \nu\rho \rangle}{\partial x} \right\} c_{\nu q}$$

$$\langle pq \,||\, rs \rangle^{x} = \sum_{\mu\nu\sigma\rho} c^{*}_{\mu p} c^{*}_{\sigma q} \frac{\partial \langle \mu\sigma \,||\, \nu\rho \rangle}{\partial x} c_{\nu r} c_{\rho s}$$

$$S^{x}_{pq} = \sum_{\mu\nu} c^{*}_{\mu p} \frac{\partial S_{\mu\nu}}{\partial x} c_{\nu q}$$

For implementation expressed in AO basis

Flowchart CC gradients



Computation independent of N_{pert}

→ for geometrical gradients: independent of number of atoms

A note on orbital relaxation

- For the CC (or MP, CI, …) wave function:

$$\Psi = \Psi(c, t)$$

MO coefficients

amplitudes (or CI coefficients, …)

switch on perturbation x

$$\Psi = \Psi(c(x), t(x))$$

MOs may react to the perturbation

or

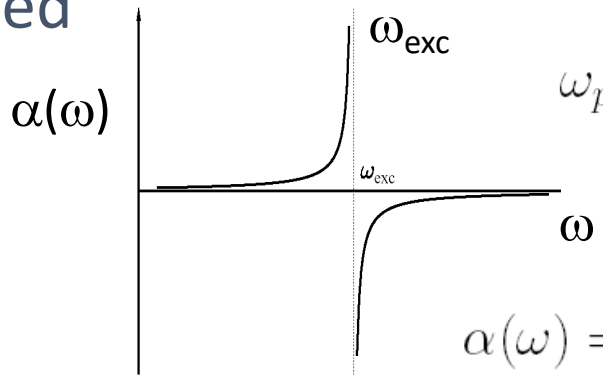
$$\Psi = \Psi(c(0), t(x))$$

Perturbation only switched on for correlation treatment
NO orbital relaxation

- Choice whether to include orbital relaxation, i.e., terms that stem from the HF wave-function parameters (Brillouin & orthogonality condition)

A note on orbital relaxation

- Relevance of inclusion depends on type of perturbation
  - Perturbation changes the metric $\frac{dS_{\mu\nu}}{dx} \neq 0$ → inclusion important
    - In particular for geometric perturbations a must (correct gradient)
  - Perturbation does not change the metric → free choice
    - Electric field
  - For frequency-dependent perturbations, consideration of orbital relaxation leads to additional unphysical poles (from HF) → not recommended



$$\omega_{pole} = (E_{exc} - E_0)/\hbar$$

$$\alpha(\omega) = \sum_i \frac{|\langle\Psi_0|\hat{\mu}|\Psi_i\rangle|^2}{\omega_{oi} - \omega} + \sum_i \frac{|\langle\Psi_0|\hat{\mu}|\Psi_i\rangle|^2}{\omega_{oi} + \omega}$$

⇒ pole position at excitation frequency

Take-home messages so far

- Properties can be calculated as energy derivatives
- Analytic differentiation is preferred (accuracy + cost)
- Perturbations that change the metric can cause more work (validity of Hellman-Feynman theorem)
- Method of Lagrange multipliers is useful to get (efficient) expressions for the derivatives also for non-variational methods
- Usefulness of inclusion of orbital relaxation depends on property