Molecular structures

Minimization, vibrational analysis, transition states

Marcus Elstner, Tomáš Kubař & Christoph Jacob

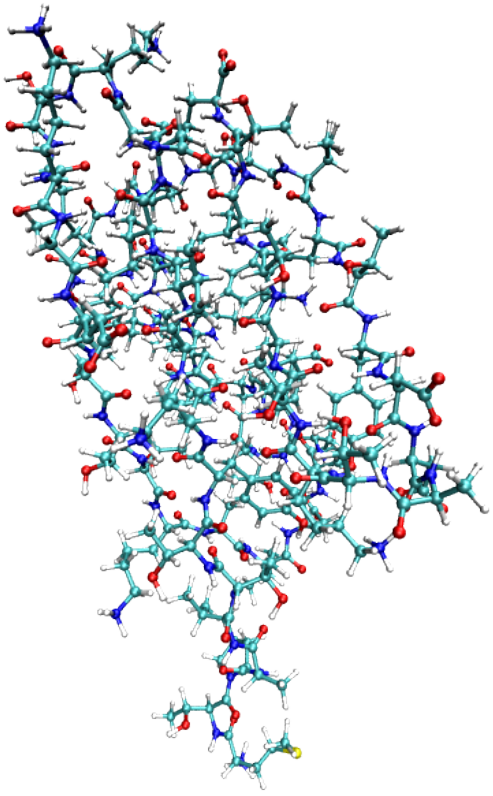
2018, June 26

Modeling of biomolecules

Potential energy surface: $E_{\mathrm{el}} = E_{\mathrm{el}}(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)$
(coordinates of atoms/nuclei $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N$)

Approximations:

- Born–Oppenheimer approximation
  - separation of nuclei and electrons
  - E_{el} obtained for fixed positions of nuclei
- Classical description of nuclei
  - rather than quantum mechanics for the motion of the nuclei
- Application of a force field
  - harmonic springs
  - point-charge electrostatics
  - ...



Energy with a force field

$$E_{\text{el}}(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N) = \frac{1}{2} \sum_{i}^{\text{bonds}} k_i (r_i - r_i^0)^2 + \frac{1}{2} \sum_{i}^{\text{angles}} k_j^{\theta} (\theta_j - \theta_j^0)^2$$

$$+ \frac{1}{2} \sum_{n}^{\text{torsions}} \cos(n_n \omega_n - \gamma_n)$$

$$+ \sum_{I<J} \frac{1}{4\pi\varepsilon_0} \frac{q_I q_J}{R_{IJ}} + \sum_{I<J} 4\varepsilon_{IJ} \left(\left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{12} - \left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{6} \right)$$

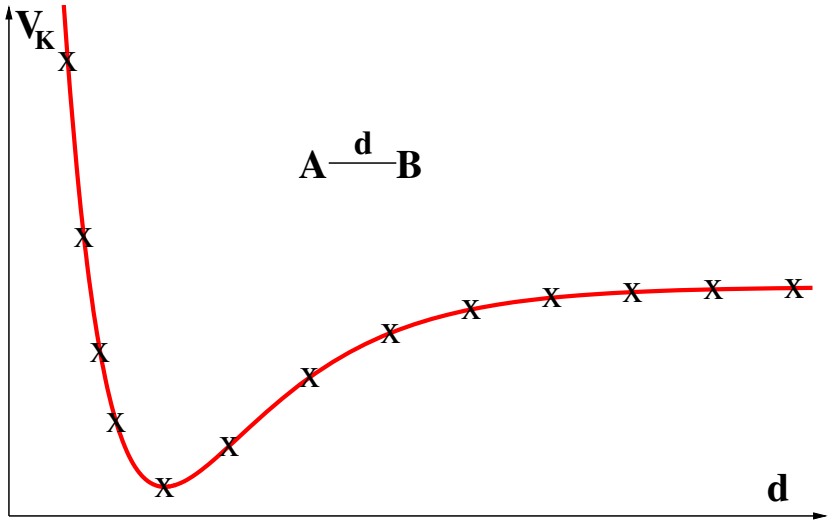
Potential energy surface

electronic energy is a function of coordinates of nuclei $\{\boldsymbol{R}_I\}$

$$E = E(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)$$

$\rightarrow$ electronic energy defines the potential energy surface (PES)

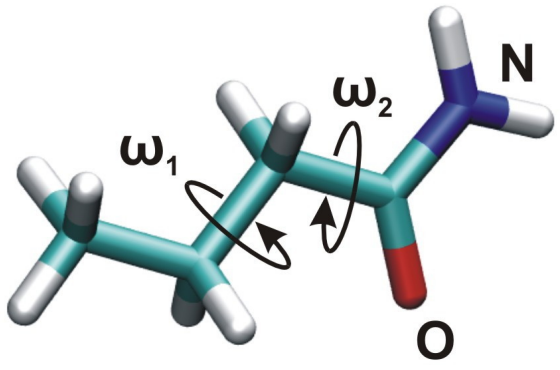
example: a diatomic molecule

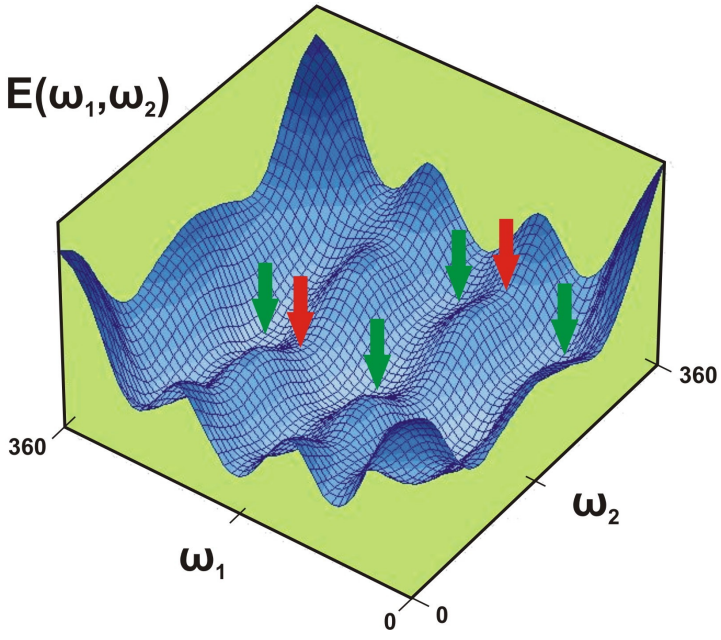


$\rightarrow$ all of the calculations only provide the PES point-wise

Sampling of the potential energy surface

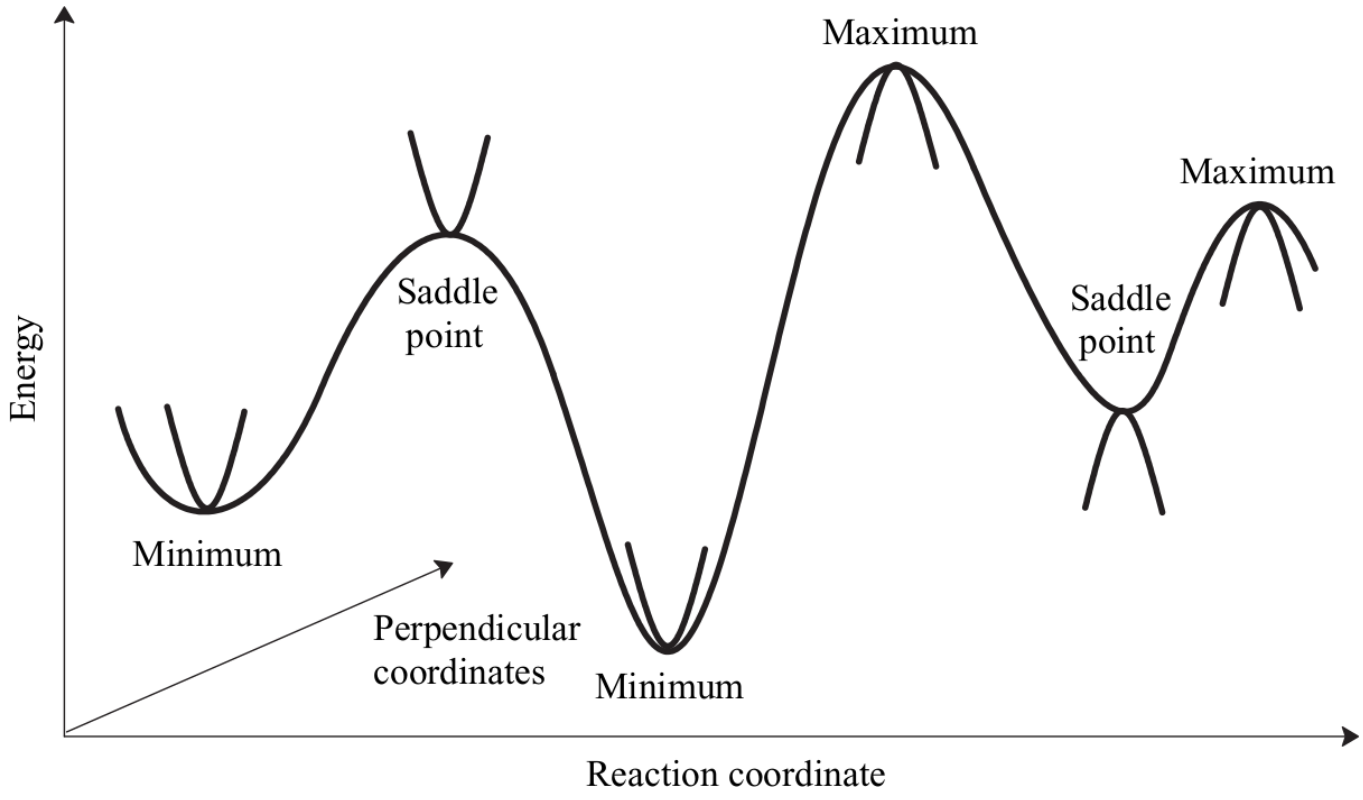
chemically interesting:
stationary points
on the PES





- minimum:
  'equilibrium' structure,
  stable conformation of a molecule
- saddle point of 1st order:
  transition state (TS),
  point of maximum energy along the direction of a 'reaction'

Potential energy surface

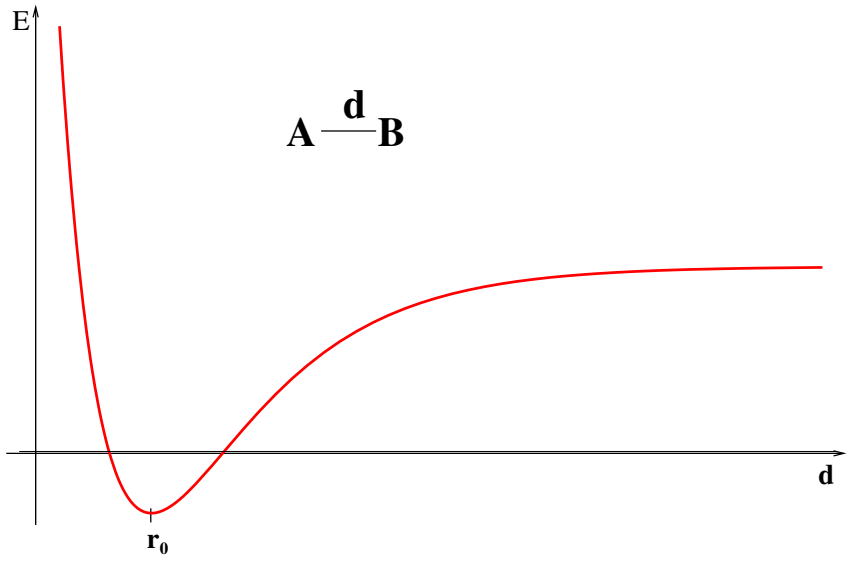


- **minimum**: any change of structure $\rightarrow$ increase of energy
- **saddle point of 1st order**: maximum along one coordinate, (reaction coordinate) $\times$ minimum along any other coordinate

Characterization of stationary points

How shall we find the interesting stationary points?

Example: Minimum of energy of a diatomic molecules



- condition for a stationary points: $\frac{\mathrm{d}E}{\mathrm{d}d} = 0$
- additional condition for a minimum: $\frac{\mathrm{d}^2E}{\mathrm{d}d^2} \geq 0$

Characterization of stationary points

How shall we find the interesting stationary points?

generally: $3N$ atom coordinates; $E = E(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)$

Condition for stationary points:

$$\text{Gradient: } \boldsymbol{g} = \boldsymbol{\nabla} E = \left(\frac{\mathrm{d}E}{\mathrm{d}x_1}, \frac{\mathrm{d}E}{\mathrm{d}y_1}, \frac{\mathrm{d}E}{\mathrm{d}z_1}, \frac{\mathrm{d}E}{\mathrm{d}x_2}, \ldots, \frac{\mathrm{d}E}{\mathrm{d}z_N} \right)^T = 0$$

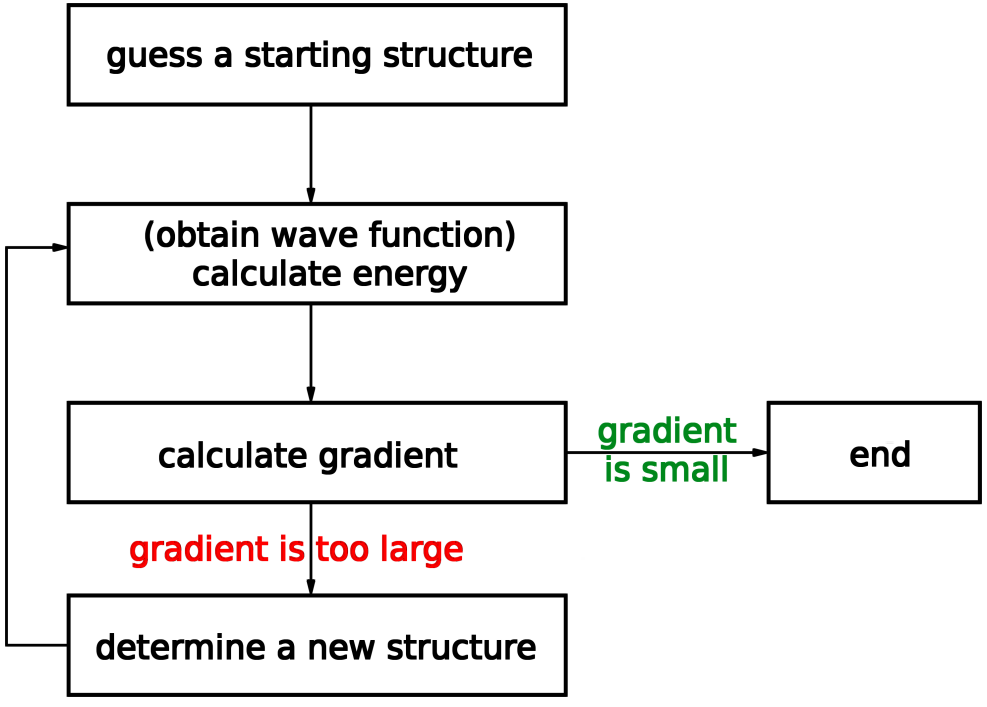
Second derivatives:

$$\text{Hessian: } \boldsymbol{H} = \begin{pmatrix} \frac{\mathrm{d}^2E}{\mathrm{d}x_1^2} & \frac{\mathrm{d}^2E}{\mathrm{d}x_1\mathrm{d}y_1} & \frac{\mathrm{d}^2E}{\mathrm{d}x_1\mathrm{d}z_1} & \frac{\mathrm{d}^2E}{\mathrm{d}x_1\mathrm{d}x_2} & \cdots \\ \frac{\mathrm{d}^2E}{\mathrm{d}y_1\mathrm{d}x_1} & \frac{\mathrm{d}^2E}{\mathrm{d}y_1^2} & \frac{\mathrm{d}^2E}{\mathrm{d}y_1\mathrm{d}z_1} & \frac{\mathrm{d}^2E}{\mathrm{d}y_1\mathrm{d}x_2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- **Condition for minimum**: all eigenvalues of $\boldsymbol{H}$ are positive
- **Saddle point of 1st order**: one eigenvalue of $\boldsymbol{H}$ is negative, all of the others are non-negative

Geometry optimization = Energy minimization

→ search for a local minimum, starting from a suitable structure



- starting structure?
  – depends on the chemical problem
- convergence criterion? e.g.

$$|\boldsymbol{g}| = \sqrt{\sum_{I,\alpha} g_{I,\alpha}^2} < t$$

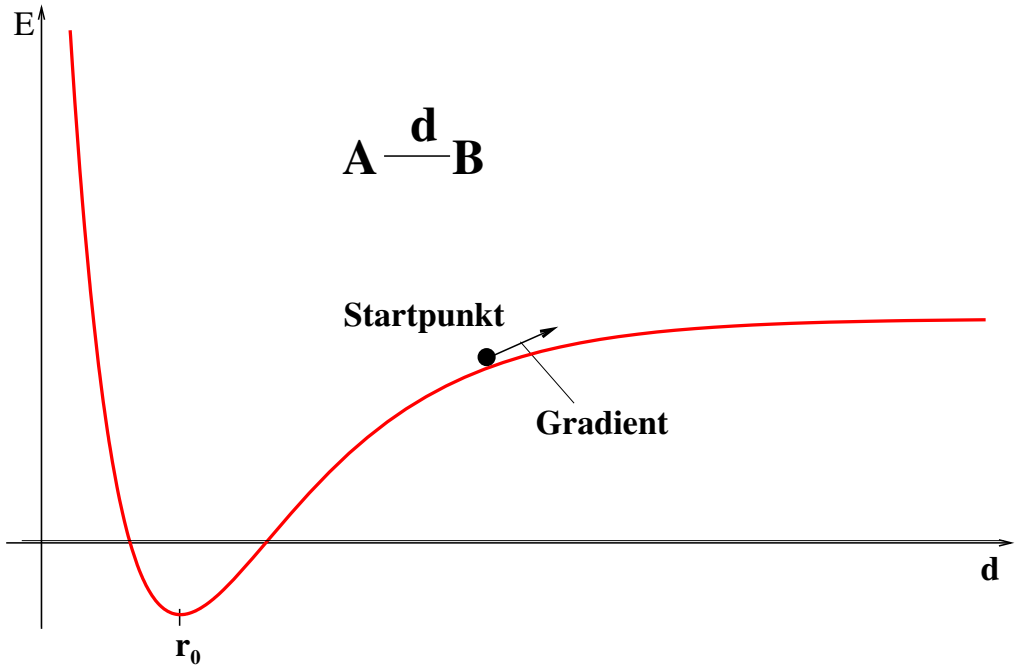
- how shall we determine the new structure?

- try to make as few steps as possible
  (calculation of energy/gradients is expensive)
- avoid any calculation of Hessian
  (that is even more expensive)

Geometry optimization = Energy minimization

How to make a step towards a minimum?

Example: geometry optimization of a diatomic molecule



- follow the negative of gradient: $\boldsymbol{R}_{i+1} = \boldsymbol{R}_i + \Delta\boldsymbol{R}_i = \boldsymbol{R}_i - \alpha\boldsymbol{g}_i$
- how shall the step length α be determined?

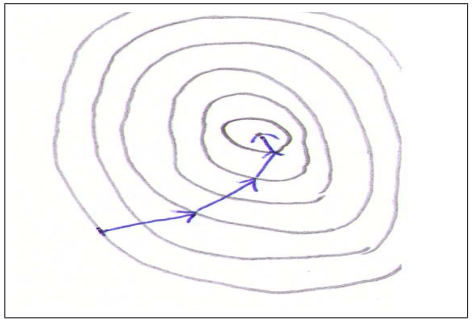
Steepest descents optimization

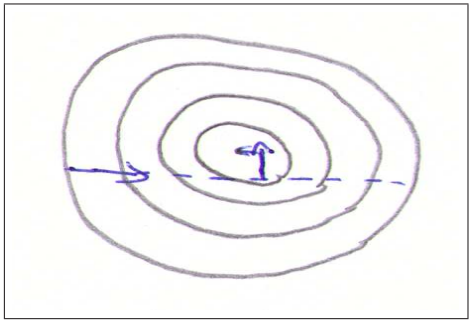
- Step along the negative of gradient

$$\Delta \boldsymbol{R}_i = \alpha \boldsymbol{d}_i \qquad \boldsymbol{d}_i = -\boldsymbol{g}_i$$

- Choice of the step length α?
  - too short $\rightarrow$ too many steps needed
  - too long $\rightarrow$ overshoot the minimum
- '*line search*':

  choose α such that the energy
  in direction of gradient $\boldsymbol{d}_i = -\boldsymbol{g}_i$
  is minimized
  – calculate the energy
  at several points along a line
- Then, the convergence is guaranteed.

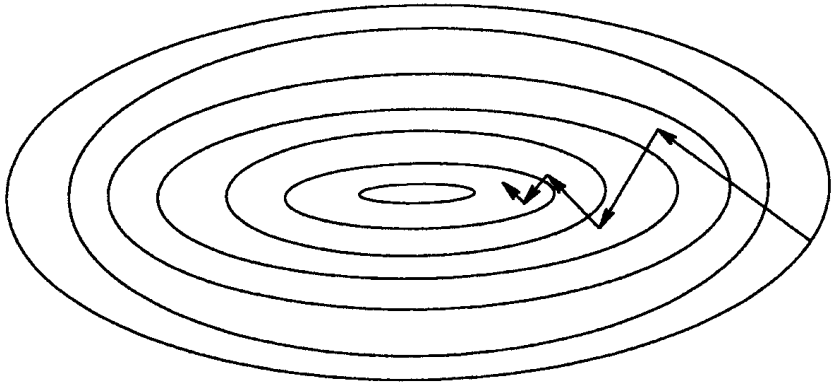




Conjugate gradient optimization

Problems of steepest descents

- too many steps in similar directions (in narrow valleys)
- convergence is getting slower when close to the minimum



Solution – conjugate gradient optimization

- make the step $\boldsymbol{d}_i$ orthogonal to *all* of the previous ones

$$\boldsymbol{d}_i = -\boldsymbol{g}_i + \beta_i \boldsymbol{d}_{i-1}$$

- various schemes: Fletcher–Reeves or Polak–Riebere:

$$\beta_i^{\mathrm{FR}} = \frac{\boldsymbol{g}_i^t \boldsymbol{g}_i}{\boldsymbol{g}_{i-1}^t \boldsymbol{g}_{i-1}} \qquad \beta_i^{\mathrm{PR}} = \frac{(\boldsymbol{g}_i - \boldsymbol{g}_{i-1})^t \boldsymbol{g}_i}{\boldsymbol{g}_{i-1}^t \boldsymbol{g}_{i-1}}$$

Newton–Raphson optimization

Taylor expansion of the PES around $\boldsymbol{R}_i$:

$$E(\boldsymbol{R}) = E(\boldsymbol{R}_i) + \boldsymbol{g}_i \cdot (\boldsymbol{R} - \boldsymbol{R}_i) + \tfrac{1}{2}(\boldsymbol{R} - \boldsymbol{R}_i)^t \cdot \boldsymbol{H}_i \cdot (\boldsymbol{R} - \boldsymbol{R}_i) + \cdots$$

minimum of this Taylor expansion up to the 2nd order:

$$\frac{\mathrm{d}E}{\mathrm{d}\boldsymbol{R}} = \boldsymbol{g}_i + \boldsymbol{H}_i \cdot (\boldsymbol{R} - \boldsymbol{R}_i) = 0 \rightarrow \boldsymbol{H}_i \cdot (\boldsymbol{R} - \boldsymbol{R}_i) = -\boldsymbol{g}_i$$

Newton–Raphson optimization:

calculate the step toward the minimum from that expression:

$$\Delta\boldsymbol{R}_i = -\boldsymbol{H}_i^{-1} \cdot \boldsymbol{g}_i$$

(on a harmonic PES, this would lead to the minimum directly)

Problem: the calculation of Hessian in every step is expensive

Quasi-Newton optimization

Apply an approximated Hessian rather than the exact Hessian

- start e.g. with a unit matrix ($\rightarrow$ 1 steepest-descents step)
- other starting Hessians may be better
  – full calculation (just once), or from certain simple rules
- in every step, use the gradients $\boldsymbol{g}_i$
  to improve the approximated Hessian $\tilde{\boldsymbol{H}}_i$,
  then invert it for $\Delta \boldsymbol{R}_i = -\boldsymbol{H}_i^{-1} \cdot \boldsymbol{g}_i$
- usually converges quickly and reliably, at the cost of
  storage for Hessian of $\mathcal{O}(N^2)$ and its inversion of $\mathcal{O}(N^3)$
- compare: CG is $\mathcal{O}(N)$ but converges more slowly
- a standard method in most quantum chemical / MD packages

Quasi-Newton optimization

various update algorithms are available

- Broyden–Fletcher–Goldfarb–Shanno (**BFGS**):

$$\tilde{\boldsymbol{H}}_{i+1} = \tilde{\boldsymbol{H}}_i + \frac{\Delta \boldsymbol{g}_i \otimes \Delta \boldsymbol{g}_i^t}{\Delta \boldsymbol{g}_i^t \cdot \Delta \boldsymbol{R}_i} - \frac{\tilde{\boldsymbol{H}}_i \cdot \Delta \boldsymbol{R}_i \otimes \Delta \boldsymbol{R}_i^t \cdot \tilde{\boldsymbol{H}}_i}{\Delta \boldsymbol{g}_i^t \cdot \tilde{\boldsymbol{H}}_i \cdot \Delta \boldsymbol{R}_i}$$

  - symmetric and positive-definite
  - minimizes the change in Hessian

- limited-memory BFGS (**L-BFGS**)
  - propagate the inverse Hessian (not the Hessian itself)
    $\rightarrow$ the $\mathcal{O}(N^3)$ matrix inversion is eliminated
  - do not store the Hessian or inverse Hessian in memory
    - rather, reconstruct the matrix on the fly
      from $\Delta \boldsymbol{g}_i$ & $\Delta \boldsymbol{R}_i$ over a few (< 10) last steps

    $\rightarrow$ the $\mathcal{O}(N^2)$ storage requirement is eliminated

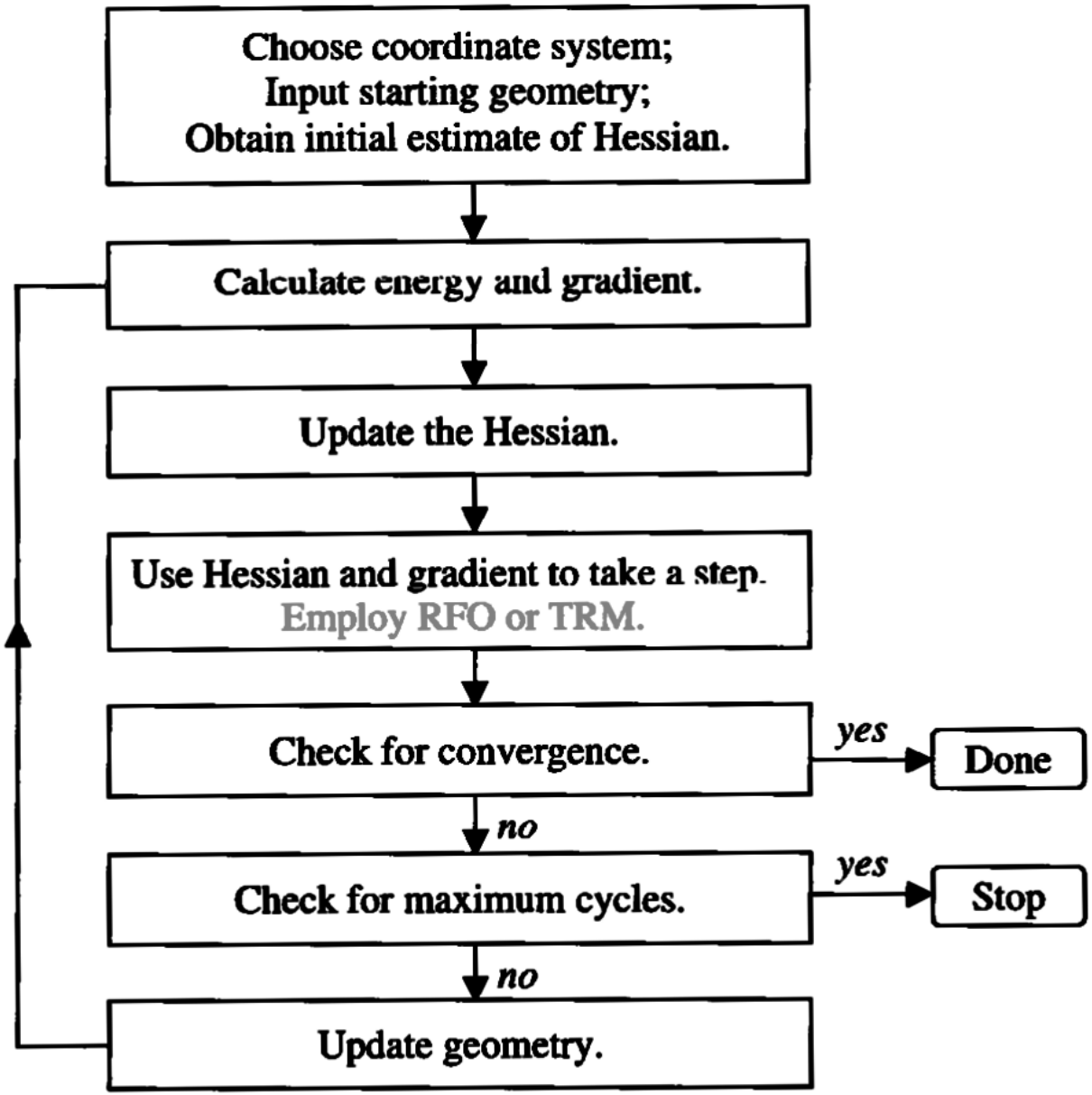
Quasi-Newton optimization

q-N converges the better, the closer PES is to a quadratic form
– depends on the choice of coordinate system strongly

Possible choices of coordinate system

- cartesian coordinates
  – simple, but not adjusted to a 'chemical' problem
  → often slow convergence
- internal coordinates
  – use bond lengths, angles and torsional angles
  → often good convergence
  → but the definition of $3N - 6$ coordinates difficult
- redundant internal coordinates
  – use 'too many' internal coordinates
  → mostly good convergence
  → simple automatized definition is possible
- more complex coordinate systems possible

Quasi-Newton optimization



Summary: geometry optimization

- starting point: a chemically meaningful structure
- minimization procedures:
  - *steepest descents:* converges always, but slowly
  - better: *conjugate gradients*, *quasi-Newton (e.g. BFGS)*
    – all of these avoid the calculation of Hessian
- biomolecules – often very difficult to find true minima
- in the quantum chemistry
  - calculations mostly limited to a single minimum
  - starting point for the calculation of properties (spectra. . .)
- with the force field methods
  - starting point for MD – steepest descents is often good enough
  - pre-optimization for quantum chemical calculations

Limitations of the force field approximation

The parameters have to be determined / fitted

- difficult for certain / unusual elements (e.g., transition metals)

Conceptual limitations

- chemical bonds cannot be broken or created
- atom types are pre-determined
- atomic charges are pre-determined and constant
  – no change of electron density can be described

Electrons are not described explicitly, so no wave function

- only ground-state energy and forces are available
- no spectroscopic properties (interaction with light)
- no photochemistry (excited states)

Quantum chemistry vs. force fields

Quantum mechanics:
electronic wave functions: $\Psi_{\text{el}} = \Psi_{\text{el}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$
(coordinates of electrons $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$)
solve the electronic Schrödinger equation

$$\hat{H}^{\{\boldsymbol{R}\}}\psi^{\{\boldsymbol{R}\}} = E_{\text{el}}(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)\psi^{\{\boldsymbol{R}\}}$$

Quantum chemistry

- wave function theory (WFT)
- density functional theory (DFT)

$\rightarrow$ calculation of energy is computationally intensive

Force field methods

- evaluate E_{el} 'directly', do not look for the electronic structure

$\rightarrow$ calculation of energy is very quick and efficient