# Molecular structures

## Minimization, vibrational analysis, transition states

Marcus Elstner, Tomáš Kubař & Christoph Jacob

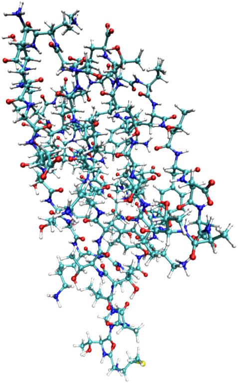
2017, June 23

# 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
  - $E_{\mathrm{el}}$ 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

$$
\begin{aligned}
E_{\mathrm{el}}(\boldsymbol{R}_1, \dots, \boldsymbol{R}_N) = {} & \frac{1}{2} \sum_{i}^{\text{bonds}} k_i (r_i - r_i^0)^2 + \frac{1}{2} \sum_{i}^{\text{angles}} k_j^{\theta} (\theta_j - \theta_j^0)^2 \\
& + \frac{1}{2} \sum_{n}^{\text{torsions}} \cos(n_n \omega_n - \gamma_n) \\
& + \sum_{I<J} \frac{1}{4\pi\varepsilon_0} \frac{q_I q_J}{R_{IJ}} + \sum_{I<J} 4\varepsilon_{IJ} \left( \left( \frac{\sigma_{IJ}}{R_{IJ}} \right)^{12} - \left( \frac{\sigma_{IJ}}{R_{IJ}} \right)^{6} \right)
\end{aligned}
$$

# 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: $\Psi_{\mathrm{el}} = \Psi_{\mathrm{el}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$
(coordinates of electrons $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$)
**solve the electronic Schrdinger equation**

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

**Quantum chemistry**

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

→ calculation of energy is computationally intensive

**Force field methods**

- evaluate $E_{\mathrm{el}}$ ‘directly’, do not look for the electronic structure

→ calculation of energy is very quick and efficient

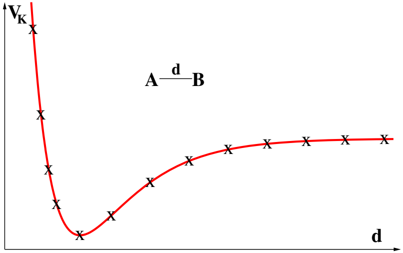
# Potential energy surface

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

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

→ electronic energy defines the potential energy surface (PES)

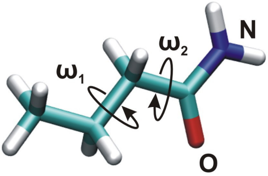
*example: a diatomic molecule*

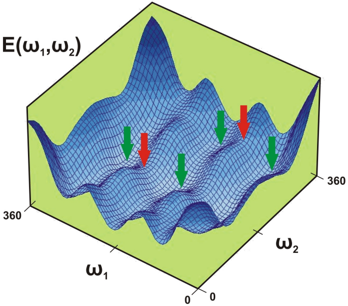


→ all of the calculations only provide the PES point-wise

# Sampling of the potential energy surface

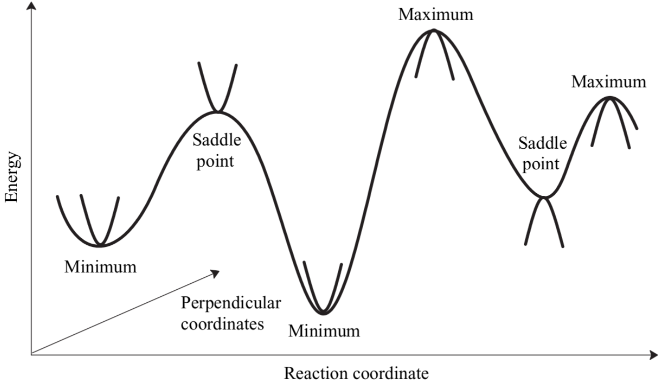
*chemically interesting*:
stationary points
on the PES





- 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

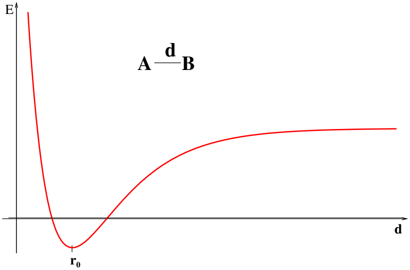


- **minimum**: any change of structure → increase of energy
- **saddle point of 1st order**: maximum along one coordinate, (reaction coordinate) × minimum along any other coordinate

# 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: $\dfrac{\mathrm{d}E}{\mathrm{d}d} = 0$
- additional condition for a minimum: $\dfrac{\mathrm{d}^2 E}{\mathrm{d}d^2} \geq 0$

# 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

# Characterization of stationary points

**How shall we find the interesting stationary points?**

- Is it possible to sample the PES systematically?
  – example: 10 atoms, $3N - 6 = 24)$ coordinates
  10 points per coordinate → $10^{24}$ calculations of energy!
  100 points per coordinate → $10^{240}$ calculations!

→ finding the global minimum – very difficult problem

- chemical intuition – concentrate on meaningful structures

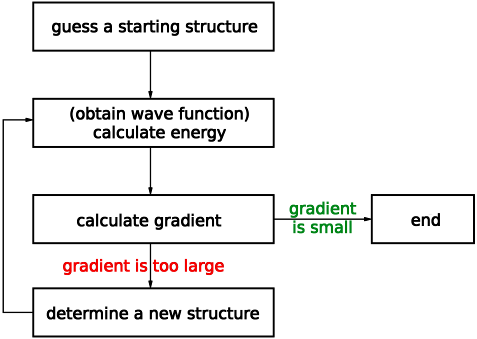
→ local minima, transition states between low-energy regions

**What do we need now?**

- calculation of gradients (and Hessian)
- algorithm
  - to optimize the geometry (search for a local minimum)
  - to search for transition states

# Geometry optimization

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



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

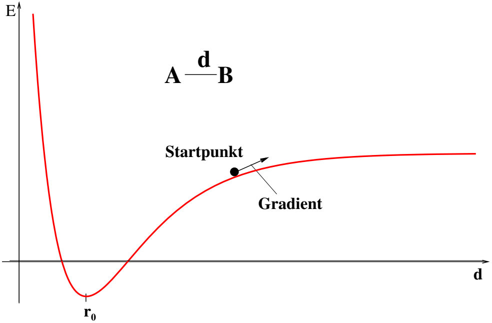
$$|\boldsymbol{g}| = \sqrt{\sum_{I,\alpha} g_{I,\alpha}^2} < t$$

- 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

**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 $\alpha$ be determined?

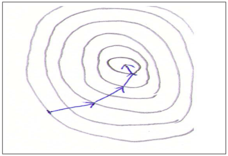
# Steepest descents optimization

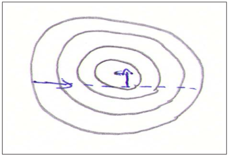
- Step along the negative of gradient

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

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

  choose $\alpha$ 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.

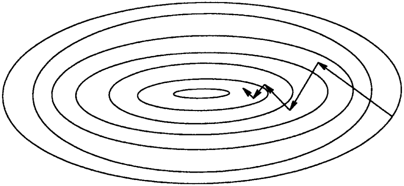




# Conjugate gradient optimization

**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 $\boldsymbol{R}_i$:**

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

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:

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

→ for harmonic PES, this leads 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 (→ steepest-descents)
- 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 $\tilde{\boldsymbol{H}}_i$
- various update algorithms are possible, e.g. **BFGS**:

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

- quasi-Newton algorithm often converges quickly and reliably
- standard method in most quantum chemical packages

# 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

# 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 around a minimum (and beyond)
  - pre-optimization for quantum chemical calculations

# Potential energy surface

**Energy is a function of atomic coordinates**

$E_{\mathrm{el},i}(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)$ – PES from quantum chemistry or force field

- Is it possible to generate the entire potential energy surface?
  – example: 10 atoms, $3N - 6 = 24$) coordinates
  10 points per coordinate → $10^{24}$ calculations of energy!
  100 points per coordinate → $10^{240}$ calculations!
- No! only concentrate on the region close to a minimum

**Search for a local minimum: geometry optimization**

- requires a chemically meaningful starting structure
- calculation of electronic energy and its gradients
  in several (few) points of the coordinate space

**In the region close to a local minimum**

- apply Taylor expansion to the PES

# Harmonic approximation

- move the origin of coordinates into the local minimum of PES:

$$\boldsymbol{R}_{I,0} = \boldsymbol{o} \quad \text{(null vector)}$$

→ the variables $\{\boldsymbol{R}_I\}$ give the deviation from minimum

- Taylor expansion of energy

$$\begin{aligned} E_{\text{el}}(\{\boldsymbol{R}_I\}) &= \underbrace{E_{\text{el}}(\{\boldsymbol{o}\})}_{const.} + \underbrace{\sum_{I=1}^{M}\sum_{\alpha=1}^{3}\left(\frac{\partial E_{\text{el}}}{\partial R_{I,\alpha}}\right)_{\boldsymbol{R}_I=\boldsymbol{o}} R_{I,\alpha}}_{\color{red}{=0}} \\ &+ \frac{1}{2}\sum_{I,J}^{M}\sum_{\alpha,\beta=1}^{3} R_{I,\alpha}\left(\frac{\partial^2 E_{\text{el}}}{\partial R_{I,\alpha}\partial R_{J,\beta}}\right)_{\substack{\boldsymbol{R}_I=\boldsymbol{o}\\ \boldsymbol{R}_J=\boldsymbol{o}}} R_{J,\beta} + \mathcal{O}(R_{I,\alpha}^3) \end{aligned}$$

→ gradients in a local minimum of PES vanish

# Harmonic approximation

- terminate the Taylor series after the second order
- write the 2nd derivatives in a matrix form – **Hessian:**

$$\boldsymbol{F} = \{F_{ij}\} = \left\{ \frac{\partial^2 E_{\text{el}}}{\partial R_i \, \partial R_j} \right\}$$

  - coupling of coordinates of different nuclei
  - the same property – covariance matrix in PCA!
- condense the atomic masses into the coordinates:

$$\begin{aligned} \boldsymbol{R}_I^{(m)} &= \sqrt{M_I} \boldsymbol{R}_I \\ \boldsymbol{F}^{(m)} &= \boldsymbol{M}^{-1/2} \cdot \boldsymbol{F} \cdot \boldsymbol{M}^{-1/2} \end{aligned}$$

- look for: a linear combination of mass-weighted coordinates that diagonalizes the Hessian → normal coordinates $\{q_i\}$

# Harmonic approximation

- normal coordinates $\{q_i\}$, normal modes $\{Q^{(i)}\}$:

$$\begin{aligned} q_i &= \sum_{j=1}^{3N} Q_j^{(i)} R_j^{(m)} \quad \text{so that} \quad \frac{\partial^2 E_{\text{el}}}{\partial q_i \, \partial q_j} = 0 \;\; \forall i \neq j \\ \boldsymbol{Q} &= (\boldsymbol{Q}^{(1)}, \cdots, \boldsymbol{Q}^{(3N)})^T \quad \text{and} \quad \boldsymbol{Q}^{(i)} = (Q_{1,x}^{(i)}, \cdots, Q_{3N,z}^{(i)}) \\ \boldsymbol{F}^{(q)} &= \boldsymbol{Q}^T \cdot \boldsymbol{F}^{(m)} \cdot \boldsymbol{Q} \quad \text{is the diagonalization problem to solve} \end{aligned}$$

- eigenvalues of energy of the harmonic oscillator
  – vibrational frequencies $\omega_i$ follow from Hessian eigenvalues

$$\omega_i = \sqrt{F_{ii}^{(q)}}$$

note: with PCA upon MD trajectory, harmonic frequencies
are obtained from the eigenvalues of the covariance matrix

- for molecules in the gas phase:
  translations+rotations are separated from the vibrations

# Summary: vibrational analysis

- starting point: local minimum on the PES
- calculate the Hessian – most computationally intensive step:

$$\boldsymbol{F} = \left\{ \frac{\partial^2 E_{\mathrm{el}}}{\partial R_i \, \partial R_j} \right\}$$

- introduce mass weighting:

$$\boldsymbol{F}^{(m)} = \boldsymbol{M}^{-1/2} \cdot \boldsymbol{F} \cdot \boldsymbol{M}^{-1/2}$$

- diagonalize the mass-weighted Hessian:

$$\boldsymbol{F}^{(q)} = \boldsymbol{Q}^{\boldsymbol{T}} \cdot \boldsymbol{F}^{(m)} \cdot \boldsymbol{Q}$$

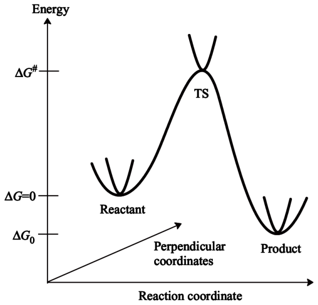
- calculate the vibrational frequencies from the eigenvalues:

$$\omega_i = \sqrt{F_{ii}^{(q)}}$$

# Transition states

How to describe a chemical reaction / another process of interest?

- structure/energy of reactant/product → reaction energy $\Delta H_r$
- structure/energy of the transition state → rate of the process



- transition state theory (TST):

$$k = \frac{k_\mathrm{B} T}{h} \cdot \exp\left[ -\frac{\Delta G^{\#}}{k_\mathrm{B} T} \right]$$

- free energy of the TS:

$$\Delta G^{\#} = \Delta H^{\#} - T \Delta S^{\#}$$

- from quantum chemistry:

$$\Delta H^{\#} = E_\mathrm{TS} - E_\mathrm{reactant}$$

- from vibrational analysis: $\Delta S^{\#}$ (via stat. thermodynamics)

# Optimization of transition states

**‘One-structure’ methods**

- start with a suitable initial structure
- optimize TS with a (quasi-)Newton method
- follow one eigenvector up
  to reach one negative eigenvalue of $\boldsymbol{H}$

Necessary:

- starting structure already close to the TS
- good starting Hessian (for quasi-Newton)
- good choice of the coordinate system
  even more important than in a usual geometry optimization

# Optimization of transition states

**'One-structure' methods**

- start with a suitable initial structure
- optimize TS with a (quasi-)Newton method
- follow one eigenvector up
  to reach one negative eigenvalue of $\boldsymbol{H}$

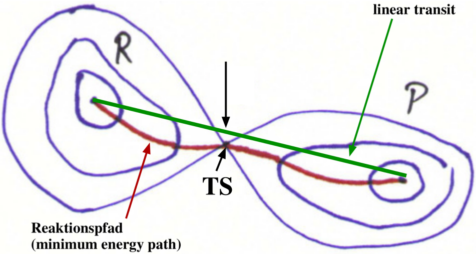
Common procedures:

- optimize the geometry of the reactant and the product
- search for a good starting structure
  - e.g. linear transit between the reactant and the product
- for quasi-Newton, usually calculate
  the full Hessian for the starting structure
  to use as a starting Hessian

# Optimization of transition states

**How to find a suitable starting structure for the TS search?**

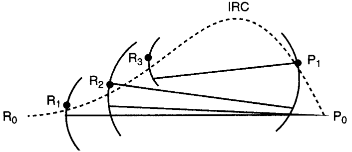
- should be good – as close as possible to the true TS
- easiest way: follow a (guessed) reaction coordinate
  e.g. bond length/angle that is changing during reaction
- linear transit: vary this coordinate in regular steps
- in every step, optimize all of the other coordinates
  ('constraint optimization' with the selected coordinate fixed)
- maximum of energy along this path → starting structure



# Optimization of transition states

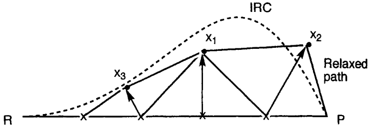
**Alternative: ‘two-structure’ methods**

- start from reactant and product; encircle the TS step-by-step



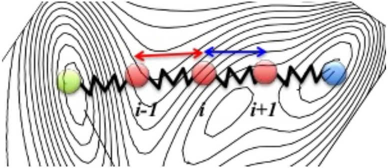
**Alternative: ‘multiple-structure’ methods**

- start with a linear interpolation between reactant and product
- optimize structuren along the entire reaction path



# Nudged elastic bands

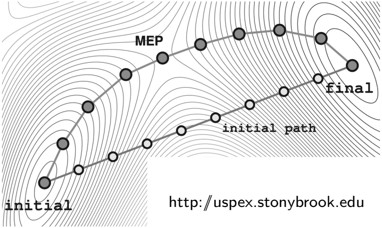
- method for finding saddle points and minimum energy paths
  between known reactants and products
- optimizes a number of intermediates along the reaction path



- each image finds the lowest energy possible
  while maintaining equal spacing to neighboring images
- restrained optimization:
  1. add harmonic springs between intermediates along the band
  2. project out the component of spring force $\perp$ to the band

# Nudged elastic bands

- construct a set of replicas of the system ($\approx 4 - 20$)
  between the initial and final state
- add a harmonic spring interaction between adjacent replicas
  – ensure continuity, much like an elastic ‘band’ would do
- during optimization, the ‘band’ induces forces on the replicas
  and brings the replicas to the minimum energy path



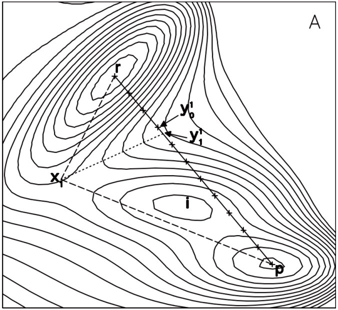
http://uspex.stonybrook.edu

# Conjugate peak refinement

CPR

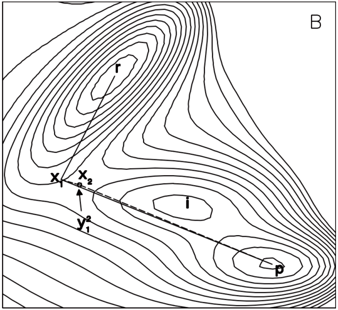
- robust method to locate TS for systems with many atoms
- iterative algorithm
- a single intermediate point is either added, refined or removed during a given iteration
- finally: the only local maxima in energy left along the path are true saddle points (TS)

# Conjugate peak refinement



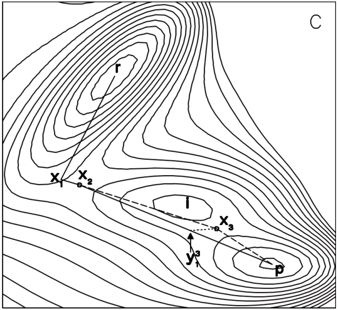
- coarse scan along the line $r - p \rightarrow$ maximum near $y_0^1$
  ($s_0$ is the direction $r - p$)
- maximization along the line $r - p \rightarrow$ maximum $y_1^1$
- minimization along conjugate vector from $y_1^1 \rightarrow$ minimum $x_1$
  (the conjugate vector is $s_1 = -g_1 + \frac{g_1^T h}{s_0^T h} s_0$ with $h \approx H s_0$)
- molecular systems – minimize along several conjug. vectors

# Conjugate peak refinement



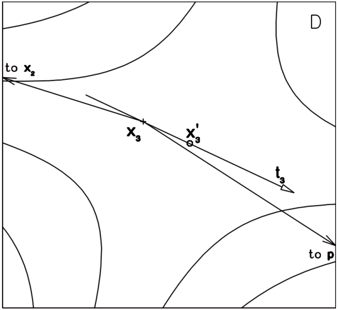
- $x_1$ is new intermediate on the path which now has 2 segments
- repeat the procedure along the lines $r - x_1$ and $x_1 - p$
- find the maximum $y_1^2$ and,
  along the conjugate direction, the minimum $x_2$
- $x_2$ is already a saddle point – meets 2 conditions:
  - RMS gradient of the energy at $x_2$ is nearly zero
  - $x_2$ is still a maximum despite resulting from line minimizations

# Conjugate peak refinement



- repeat the procedure along the line $x_2 - p$
- find the maximum $y_1^3$ and,
  along the conjugate direction, the minimum $x_3$
- RMS gradient at $x_3$ is small, so it is close
  to a stationary point (here, saddle point)

# Conjugate peak refinement



- $x_3$ is refined to produce a saddle point
- search is performed along the line $t_3$
  which is an average of $x_3 - x_2$ and $p - x_3$
- line maximization followed by conjugate line minimization(s)
  gives the saddle point $x_3'$
  - RMS gradient of the energy at $x_3'$ is nearly zero
  - $x_3'$ is still a maximum despite resulting from line minimizations

# Summary: transition states

- Optimization of TS is much more difficult
  than a normal optimization of an energy minimum
- *one-structure methods*:
  good starting structure and starting Hessian are essential
- *multiple-structure methods*:
  usually more robust, but not always available/applicable
- after the optimization of TS:
  always calculate the Hessian to check if it really is a TS!
- in quantum chemistry:
  - description of chemical reactions (→ activation energy)
  - but, reaction rates require very accurate energies
- with force field methods:
  - chemical reactions impossible, only conformational changes
  - then, the meaning of TS not always clear