C7790 Introduction to Molecular Modelling

TSM Modelling Molecular Structures

C9087 Computational Chemistry for Structural Biology

Lesson 14

Potential Energy Surface II

JS/2022 Present Form of Teaching: Rev2

Petr Kulhánek

kulhanek@chemi.muni.cz

National Centre for Biomolecular Research, Faculty of Science
Masaryk University, Kamenice 5, CZ-62500 Brno

Context

macroworld **microworld**

states

(thermodynamic properties, G, T,…)

phenomenological thermodynamics

equilibrium (equilibrium constant)

kinetics (rate constant)

free energy

(Gibbs/Helmholtz)

partition function

statistical thermodynamics

microstates

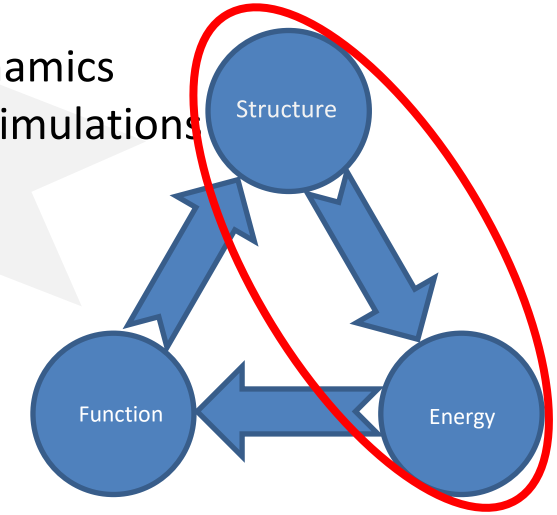
(mechanical properties, E)

Description levels (model chemistry):

- quantum mechanics
  - semiempirical methods
  - ab initio methods
  - post-HF methods
  - DFT methods
- molecular mechanics
- coarse-grained mechanics

Simulations:

- molecular dynamics
- Monte Carlo simulations
- docking
- …



microstate ≠ microworld

PES

Potential Energy Surface

Properties

Visualization

Important Points (Stationary States)

Configuration space

$$\boxed{E(\mathbf{R})}$$

R = point in 3N dimensional space (N is the number of atoms)

$$\mathbf{R} = (\underline{x_1, y_1, z_1}, x_2, y_2, z_2, \ldots, x_N, y_N, z_N)$$

Cartesian coordinates of the first nucleus (atom, bead).

individual coordinates are **degrees of freedom** of the system

The individual points form a configuration space. **Every point** in the configuration space then represents **unique structure** of the system.

Potential energy calculation

The calculation of the potential energy E(R) is possible by:

- approximate solution of the Schrödinger equation (quantum mechanics, QM)
- using empirical force fields (molecular mechanics, MM)
- hybrid QM/MM approach
- using coarse grained models

Potential energy calculation

The calculation of the potential energy E(R) is possible by:

- approximate solution of the Schrödinger equation (quantum mechanics, QM)
- using empirical force fields (molecular mechanics, MM)
- hybrid QM/MM approach
- using coarse grained models

overview of method categories

Potential energy calculation

The calculation of the potential energy E(R) is possible by:

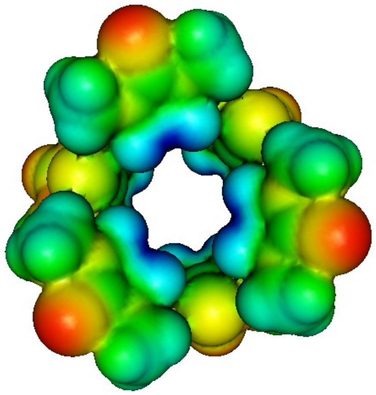
- approximate solution of the Schrödinger equation (quantum mechanics, QM)
  - HF method
  - post HF methods (MPn, CI, CC)
  - DFT methods (various functionals)
- using empirical force fields (molecular mechanics, MM)
  - forms and parameters of force fields
- hybrid QM/MM approach
  - interface, type of QM-MM interaction, link atoms, ...
- using coarse grained models

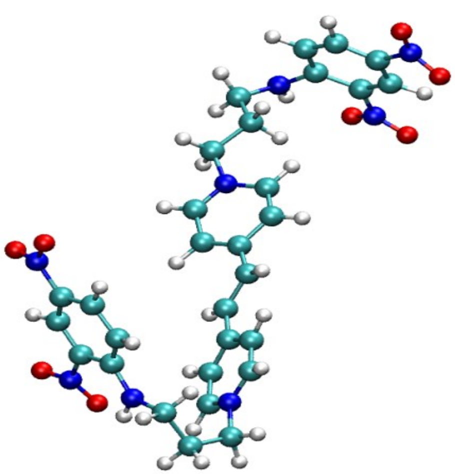
hundreds of methods differing in the approximations used

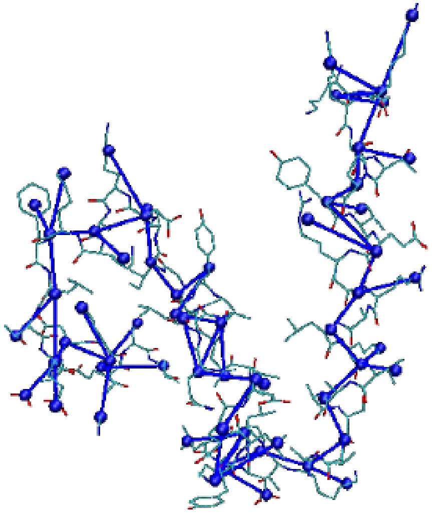
they do not affect general laws/properties of E(R)

$$\boxed{E(\mathbf{R})}$$

Potential energy calculation







QM (Quantum mechanics) | **MM (Molecular mechanics)** | ***CGM (Coarse-grained*** **mechanics)**

$$\boxed{E(\mathbf{R})} \xrightarrow{\text{approximations}} \boxed{E(\mathbf{R})} \xrightarrow{\text{approximations}} \boxed{E(\mathbf{R})}$$

R - position of atom nuclei

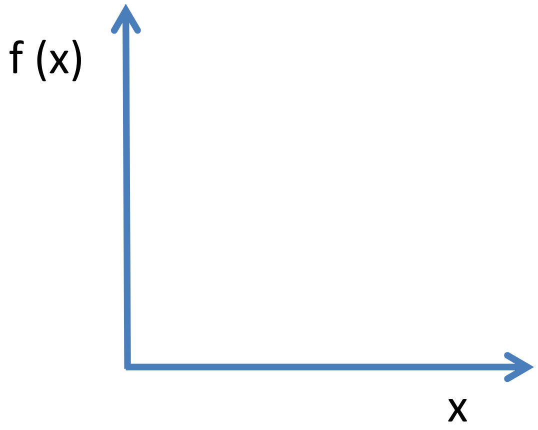
R - position of atoms

R - position of beads

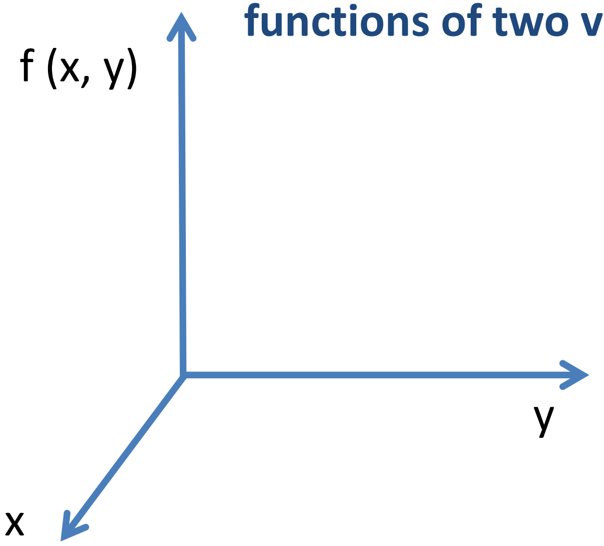
Potential energy surface can be calculated by various method (model chemistry)!

Graphical representation of functions

function of one variable

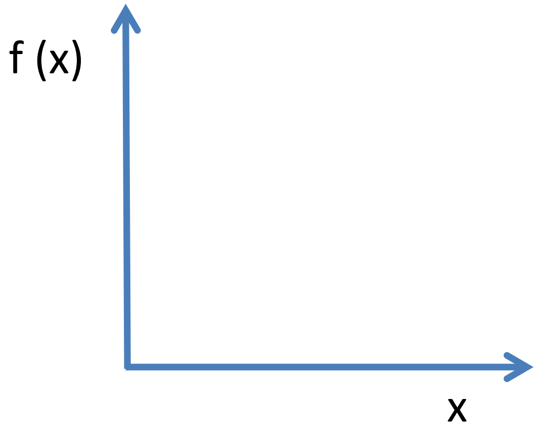


functions of two variables

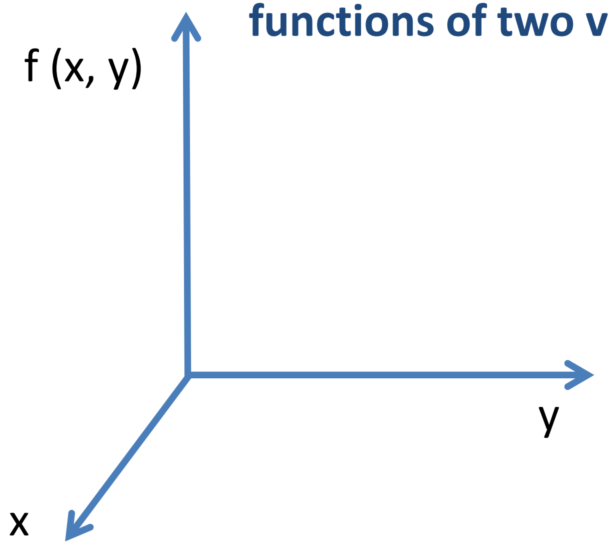


Graphical representation of functions

function of one variable

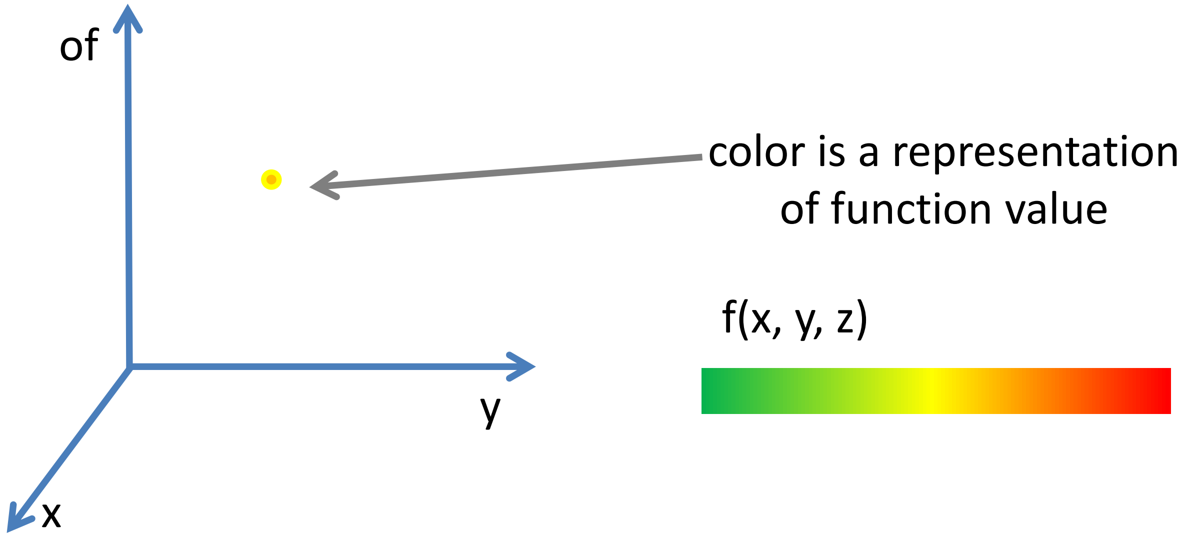


functions of two variables



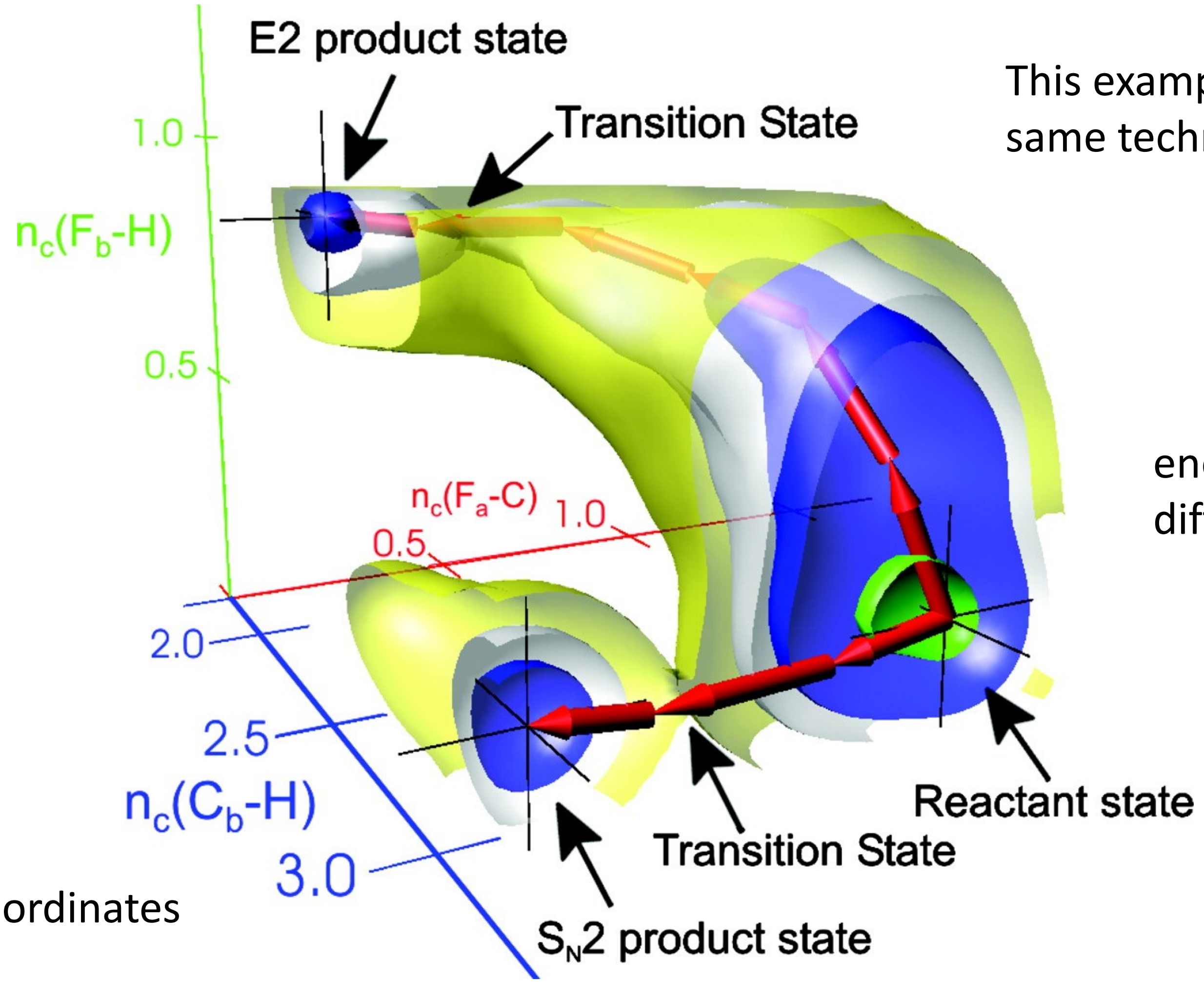
functions of three variables

volumetric display



color is a representation
of function value

Volumetric representation of functions



This example is about FES, but the same technique can be used for PES.

energy as iso surfaces with different colors

coordinates

Ensing, B.; Klein, M. L. Perspective on the Reactions between F– and CH3CH2F: The Free Energy Landscape of the E2 and SN2 Reaction Channels. *PNAS* **2005**, *102* (19), 6755–6759. https://doi.org/10.1073/pnas.0408094102.

Displaying E(R)

1 atom

$$E(x_1, y_1, z_1)$$

only volumetrically

2 atoms

$$E(x_1, y_1, z_1, x_2, y_2, z_2)$$

cannot be displayed

N atoms

$$E(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_N, y_N, z_N)$$

Example:
enzyme BsoBI has ~ 10000 atoms => 30000 degrees of freedom,
it would be necessary to use 30000 + 1 dimensional space for visualization

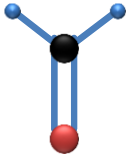
Property of E(R)

The potential energy is invariant to:

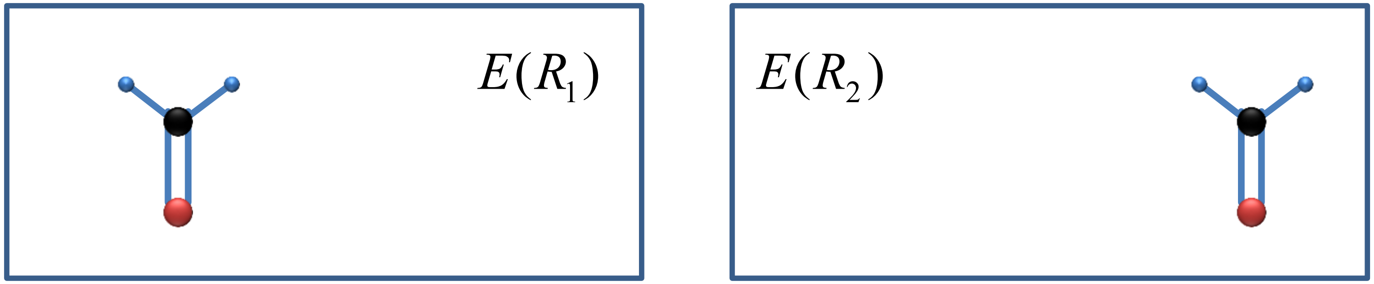
- **displacement** (translation) of entire system
- **rotation** (rotation) of entire system

} **without the action of external force fields**
(e.g., electrostatic, magnetic, etc.)

Invariance to displacement



HCHO



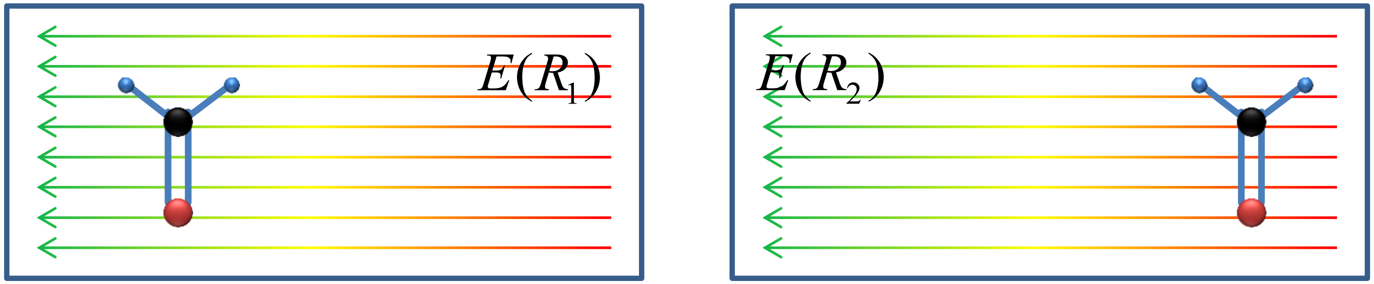
$$E(R_1) = E(R_2)$$

$$R_1 + T = R_2$$

$$T = \{\underline{x_T, y_T, z_T}, x_T, y_T, z_T, ...\}$$

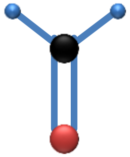
translation vector

!!! Does not apply to shift in a force field !!!

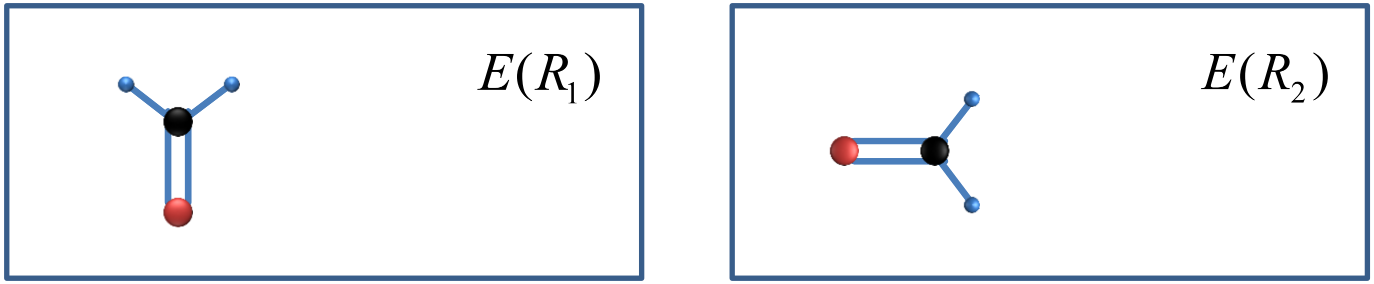


$$E(R_1) \neq E(R_2)$$

Invariance to rotation



HCHO

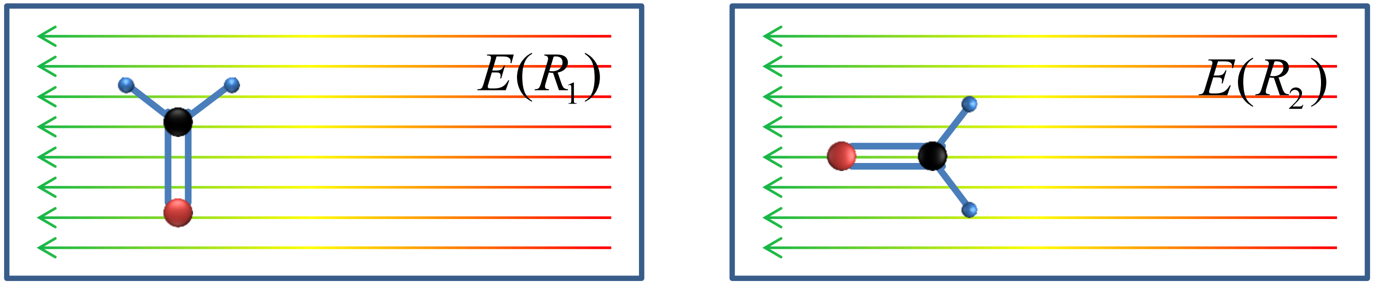


$$E(R_1) = E(R_2)$$

$$\mathbf{\Theta} R_1 = R_2$$

rotation matrix

!!! Does not apply to rotation in a force field !!!



$$E(R_1) \neq E(R_2)$$

Diatomic molecule

$$E(x_1, y_1, z_1, x_2, y_2, z_2)$$

hydrogen molecule

- three degrees of translational freedom
- two rotational degrees of freedom (molecule is linear)

Diatomic molecule

$$E(x_1, y_1, z_1, x_2, y_2, z_2)$$

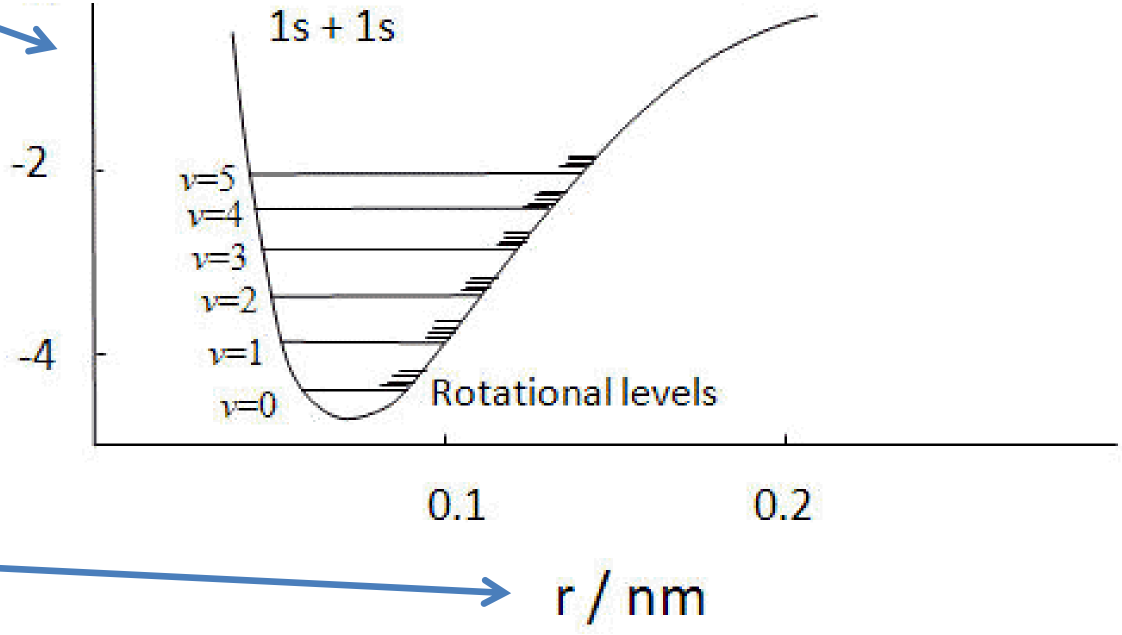
6-5 = 1

$$E(r)$$

hydrogen molecule

- three degrees of translational freedom
- two rotational degrees of freedom (molecule is linear)

hydrogen molecule

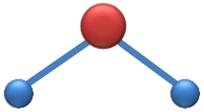


interatomic distance

Thriatomic molecule

$$E(x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3)$$

water molecule



- three degrees of translational freedom
- three rotational degrees of freedom

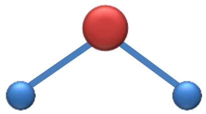
Triatomic molecule

$$E(x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3)$$

9-6=3

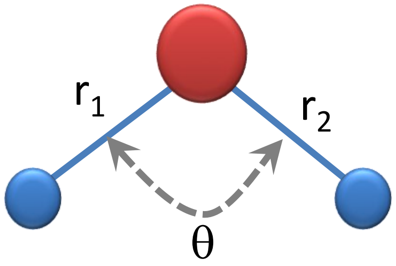
$$E(r_1, r_2, \theta)$$

water molecule



- three degrees of translational freedom
- three rotational degrees of freedom

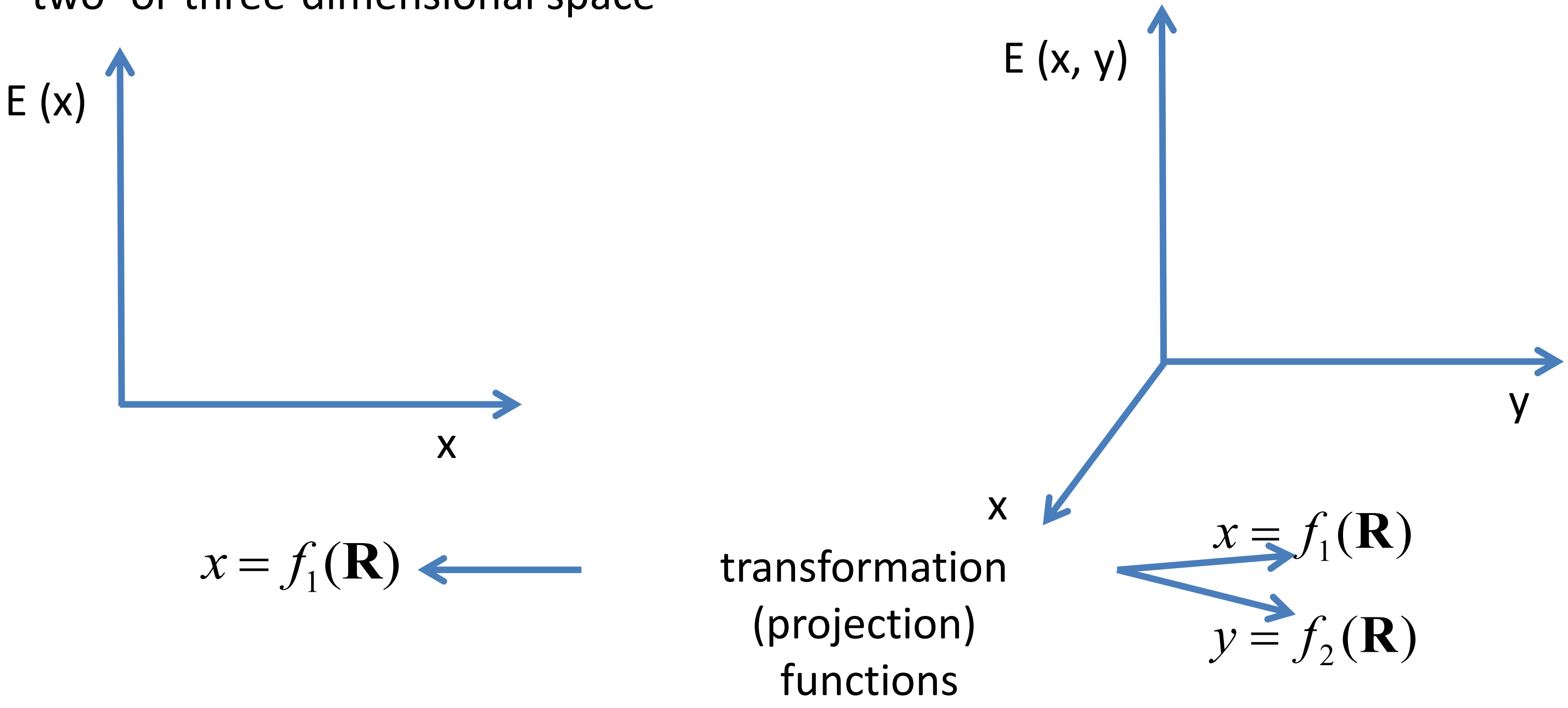
Internal coordinates r_1, r_2, θ



displayable only volumetrically

Graphical representation of E(R)

- The E(R) function is not graphically representable due to its high dimensionality.
- Thus, only its relevant part that best captures the problem being studied is displayed in a two- or three-dimensional space



$x = f_1(\mathbf{R})$

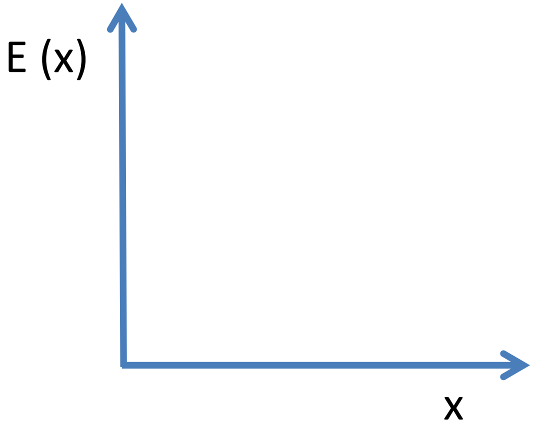
transformation (projection) functions

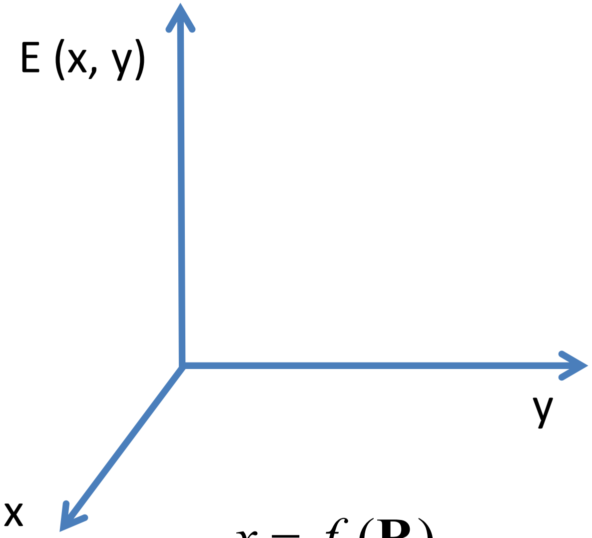
$x = f_1(\mathbf{R})$

$y = f_2(\mathbf{R})$

Graphical representation of E(R)

- The E(R) function is not graphically representable due to its high dimensionality.
- Thus, only its relevant part that best captures the problem being studied is displayed in a two- or three-dimensional space





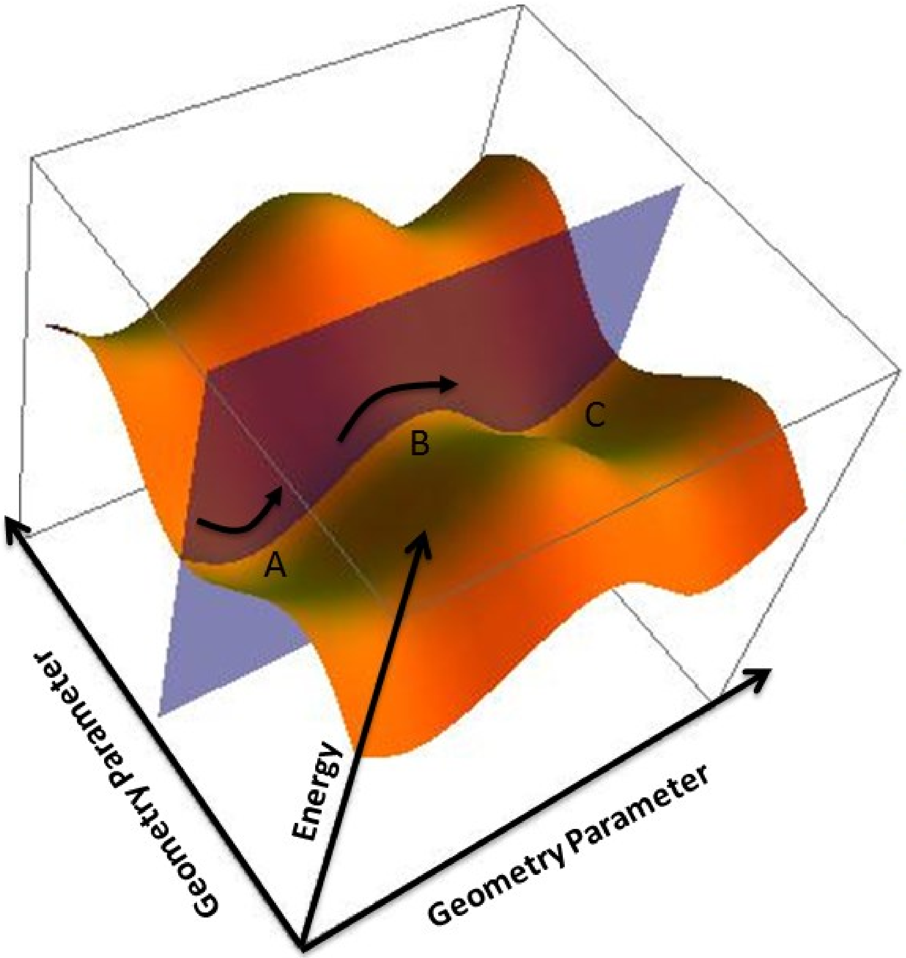
$$x = f_1(\mathbf{R}) \longleftarrow \text{transformation (projection) functions} \longrightarrow \begin{cases} x = f_1(\mathbf{R}) \\ y = f_2(\mathbf{R}) \end{cases}$$

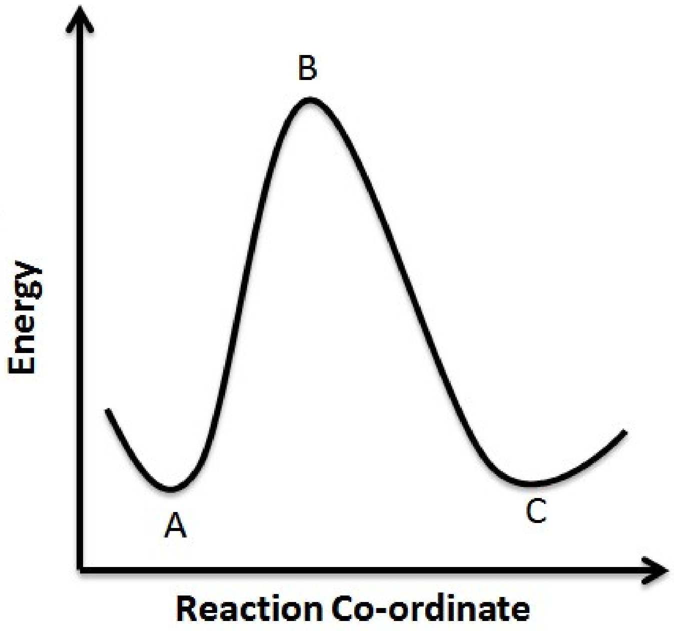
What about other degrees of freedom?

$$\mathbf{r_c} = f_2(\mathbf{R}) \longrightarrow \frac{\partial E(\mathbf{R})}{\partial \mathbf{r_c}} = 0 \longleftarrow \mathbf{r_c} = f_3(\mathbf{R})$$

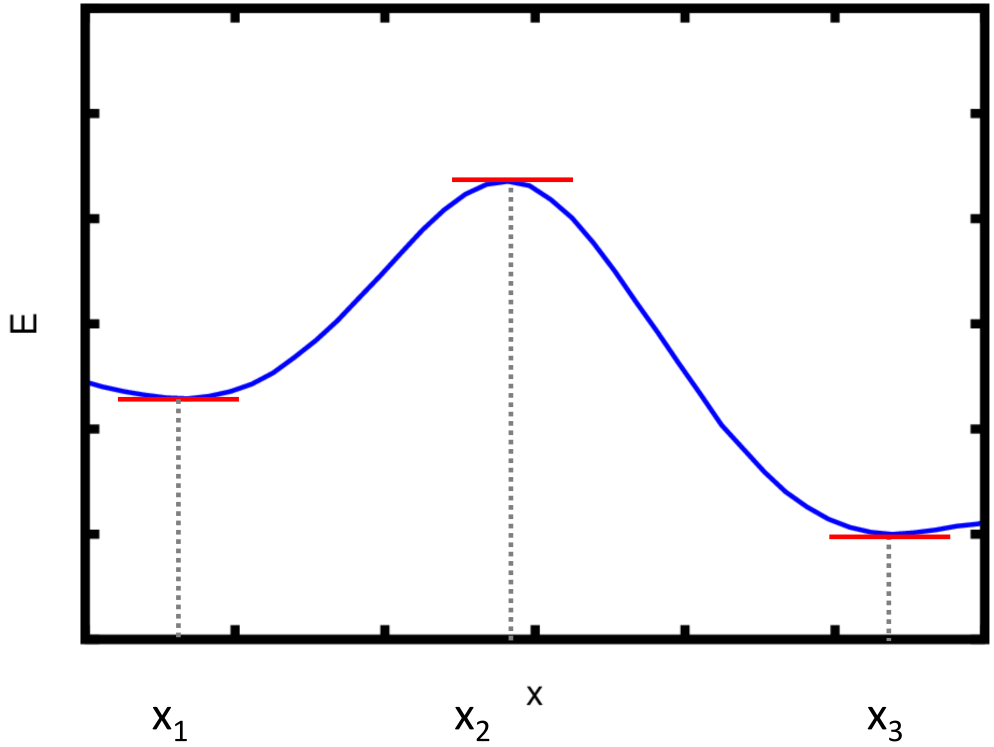
value of E(R) is minimal relative to r_C

Illustrative example





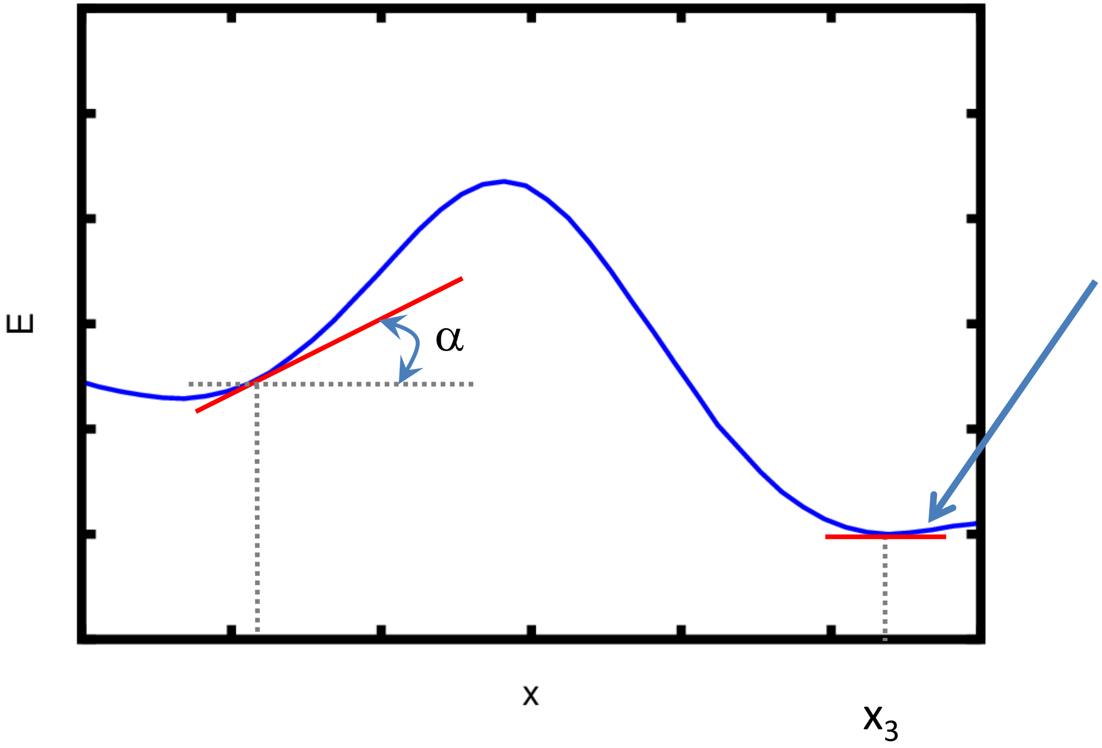
Stationary points



x_1, x_2 and x_3 are stationary points (local extrema of the function). The properties of the quantum states of the system are derived from them.

$$E_k = \boxed{E(x_1)} + E_{VRT,l}$$

Stationary points



tangent of the function E(x) at the stationary point has zero slope (α=0)

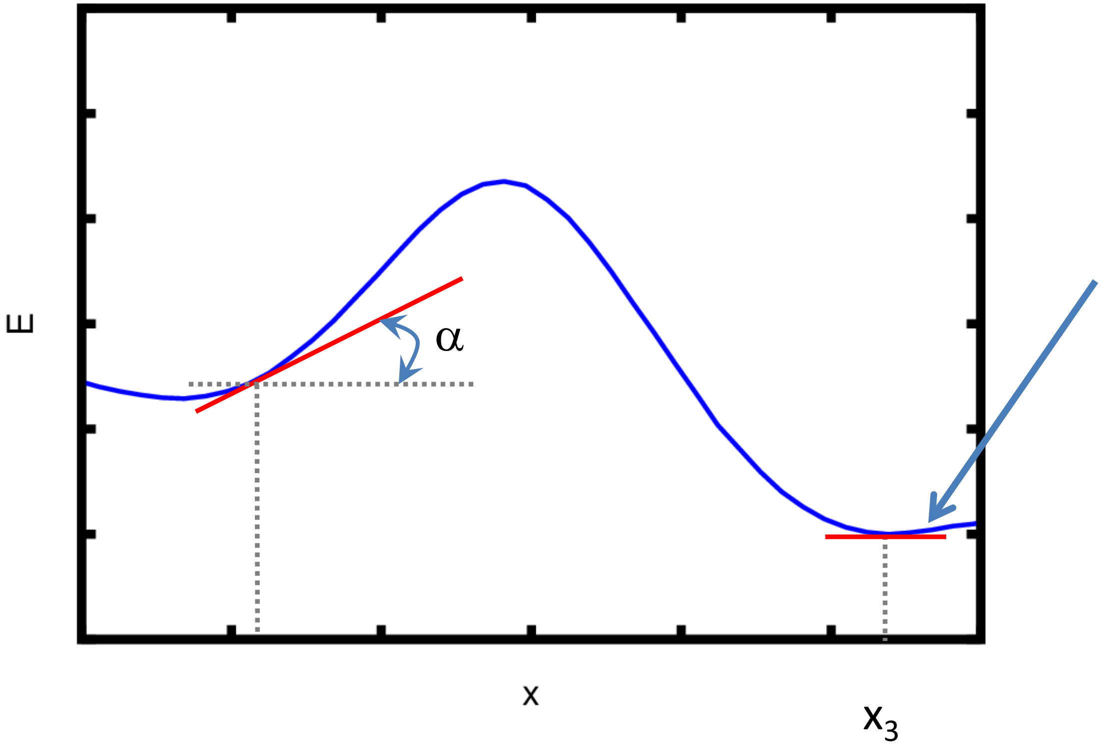
Function slope is given by **gradient of function** (i.e., first derivative of the function)

$$\tan(\alpha) = \frac{\partial E(x)}{\partial x}$$

Condition necessary for a stationary point

$$\frac{\partial E(x)}{\partial x} = 0$$

Stationary points



tangent of the function E(x) at the stationary point has zero slope (α=0)

$$\left.\frac{\partial E(x)}{\partial x}\right|_{x_3} = 0$$

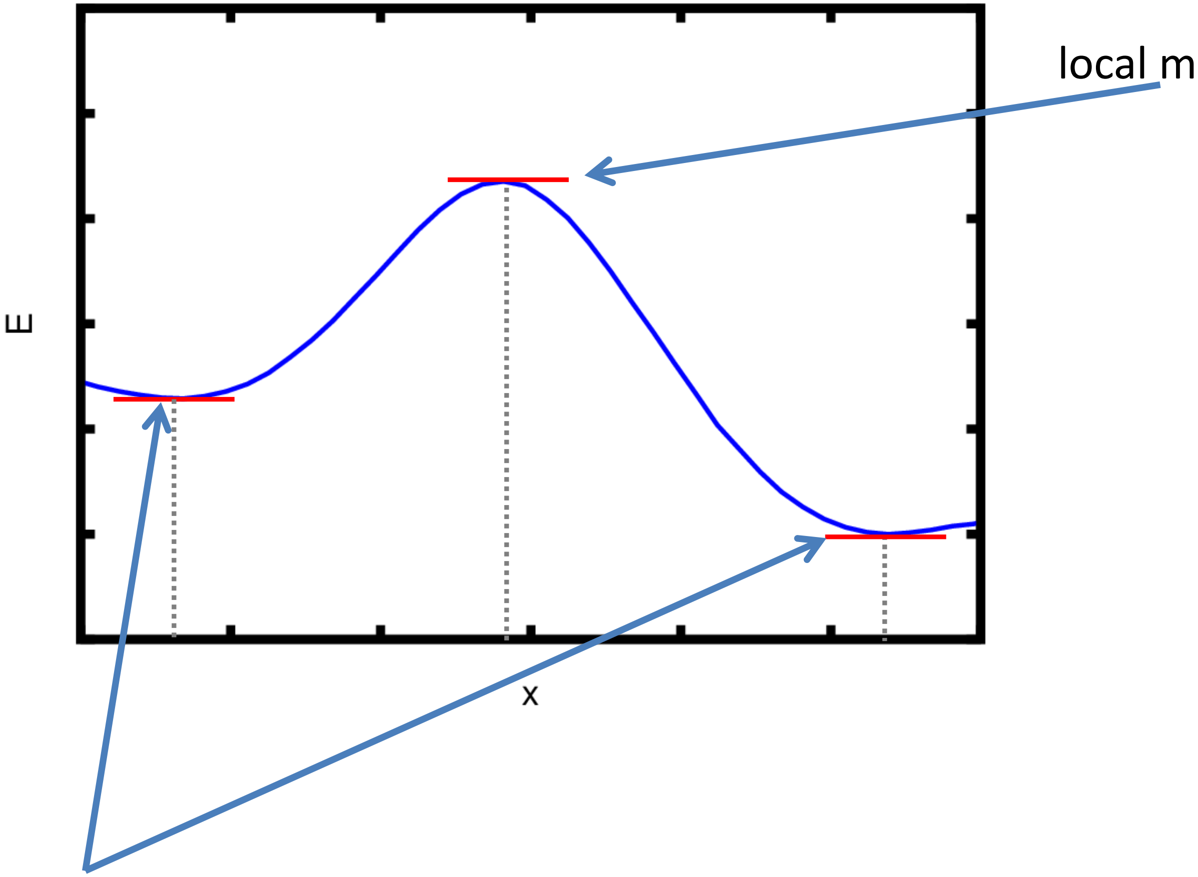
Function slope is given by **gradient of function** (i.e., first derivative of the function)

$$\tan(\alpha) = \frac{\partial E(x)}{\partial x}$$

Condition necessary for a stationary point

$$\frac{\partial E(x)}{\partial x} = 0$$

Types of stationary points



local maximum

local minima

Determining stationary point type

Taylor series:

$$E(x+\Delta x)=E(x)+\frac{\partial E(x)}{\partial x}\Delta x+\frac{1}{2}\frac{\partial^2 E(x)}{\partial x^2}(\Delta x)^2+\ldots$$

Determining stationary point type

Taylor series:

$$E(x+\Delta x)=E(x)+\frac{\partial E(x)}{\partial x}\Delta x+\frac{1}{2}\frac{\partial^2 E(x)}{\partial x^2}(\Delta x)^2+\ldots$$

gradient is zero

At the stationary point:

$$E(x+\Delta x)=E(x)+\frac{1}{2}\frac{\partial^2 E(x)}{\partial x^2}(\Delta x)^2+\ldots$$

square of the deviation is always positive

Determining stationary point type

Taylor series:

$$E(x+\Delta x)=E(x)+\frac{\partial E(x)}{\partial x}\Delta x+\frac{1}{2}\frac{\partial^2 E(x)}{\partial x^2}(\Delta x)^2+\ldots$$

gradient is zero

At the stationary point:

$$E(x+\Delta x)=E(x)+\frac{1}{2}\frac{\partial^2 E(x)}{\partial x^2}(\Delta x)^2+\ldots$$

square of the deviation is always positive

The value of function increases when deviating from a stationary point, if any **second derivative at a given point has a positive value**. The stationary point is then **local minimum**.

$$\frac{\partial^2 E(x)}{\partial x^2}>0$$

The value of function decreases when deviating from a stationary point, if any **second derivative at a given point has a negative value**. The stationary point is then **local maximum**.

$$\frac{\partial^2 E(x)}{\partial x^2}<0$$

Stationary points

Local minimum:

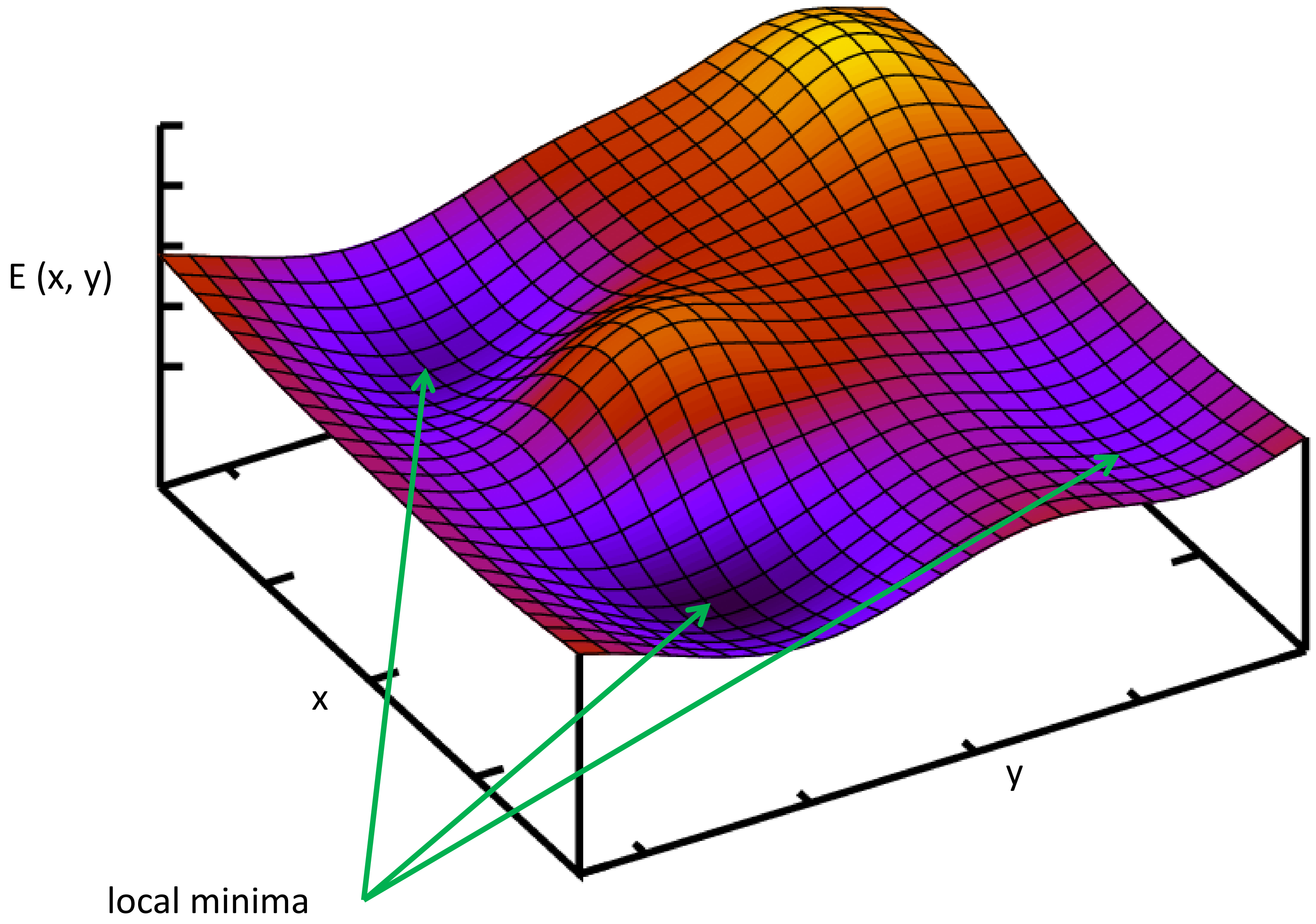
$$\boxed{\frac{\partial E(x)}{\partial x} = 0 \qquad \frac{\partial^2 E(x)}{\partial x^2} > 0}$$

Local maximum:

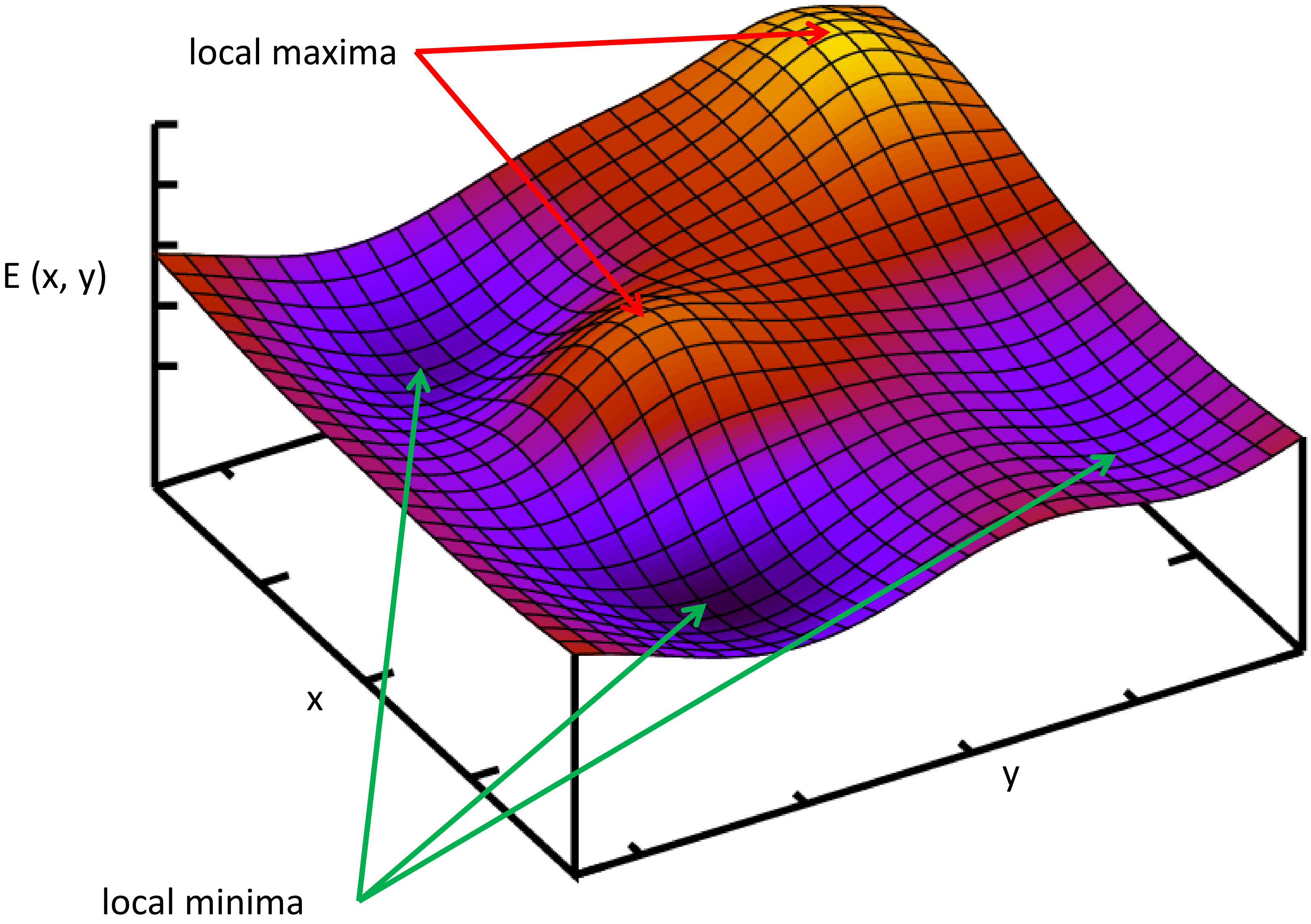
$$\boxed{\frac{\partial E(x)}{\partial x} = 0 \qquad \frac{\partial^2 E(x)}{\partial x^2} < 0}$$

!!! required condition !!!

Two-dimensional case



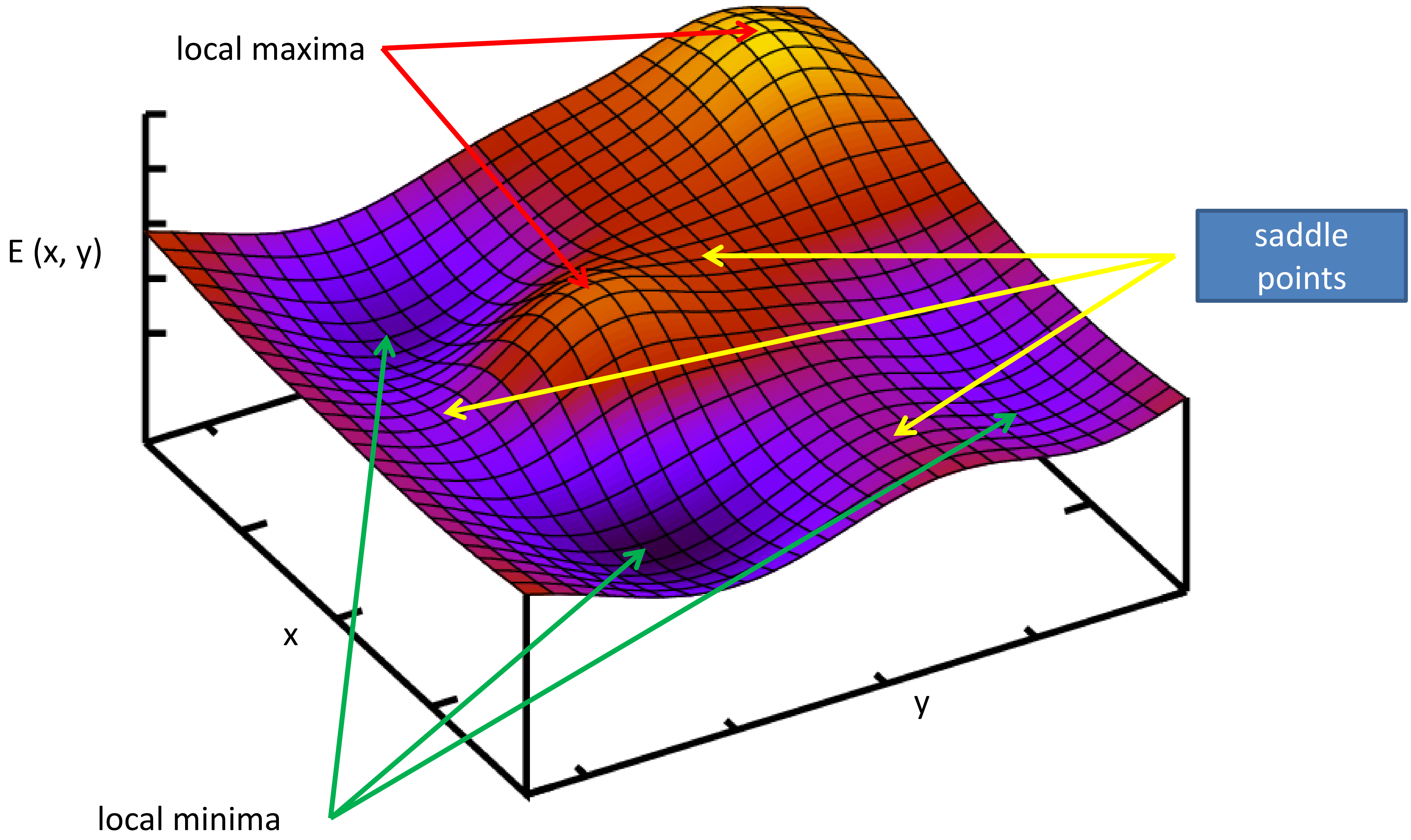
Two-dimensional case



local maxima

local minima

Two-dimensional case



Generalization for E(R)

Stationary point:

$$\frac{\partial E(\mathbf{R})}{\partial \mathbf{R}} = 0$$

required condition, each component of the gradient must be zero

gradient has 3N components

Stationary point type:

$$\begin{pmatrix}
\frac{\partial^2 E(R)}{\partial x_1^2} & \frac{\partial^2 E(R)}{\partial x_1 \partial y_1} & \frac{\partial^2 E(R)}{\partial x_1 \partial z_1} & \cdots & \frac{\partial^2 E(R)}{\partial x_1 \partial z_N} \\
\frac{\partial^2 E(R)}{\partial y_1 \partial x_1} & \frac{\partial^2 E(R)}{\partial y_1^2} & \frac{\partial^2 E(R)}{\partial y_1 \partial z_1} & \cdots & \vdots \\
\frac{\partial^2 E(R)}{\partial z_1 \partial x_1} & \frac{\partial^2 E(R)}{\partial z_1 \partial y_1} & \frac{\partial^2 E(R)}{\partial z_1^2} & \cdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 E(R)}{\partial z_N \partial x_1} & \cdots & \cdots & \cdots & \frac{\partial^2 E(R)}{\partial z_N^2}
\end{pmatrix} = \mathbf{H}(\mathbf{R})$$

Character of the stationary point is determined by **Hessian**, which is a matrix of second derivatives of potential energy.

Not to be confused with Hamiltonian!!!

N number of atoms

Properties of Hessian

$$\mathbf{H}\mathbf{c}_k = \lambda_k \mathbf{c}_k \qquad k = 1,\ldots,3N$$

eigenvalue

eigenvector

Diagonalization of Hessian is a method for finding eigenvalues and eigenvectors.

- 6 (5) eigenvalues are zero - it corresponds to the translation and rotation of the system
- remaining eigenvalues:
  - **all positive** - **local minimum**
  - **one negative, other positive** - **first order saddle point**
  - two negative, other positive - saddle point of the second order
  - .....
  - all negative - local maximum

N number of ator

Properties of Hessian

$$\mathbf{H}\mathbf{c}_k = \lambda_k \mathbf{c}_k \qquad k = 1,\ldots,3N$$

eigenvalue

eigenvector

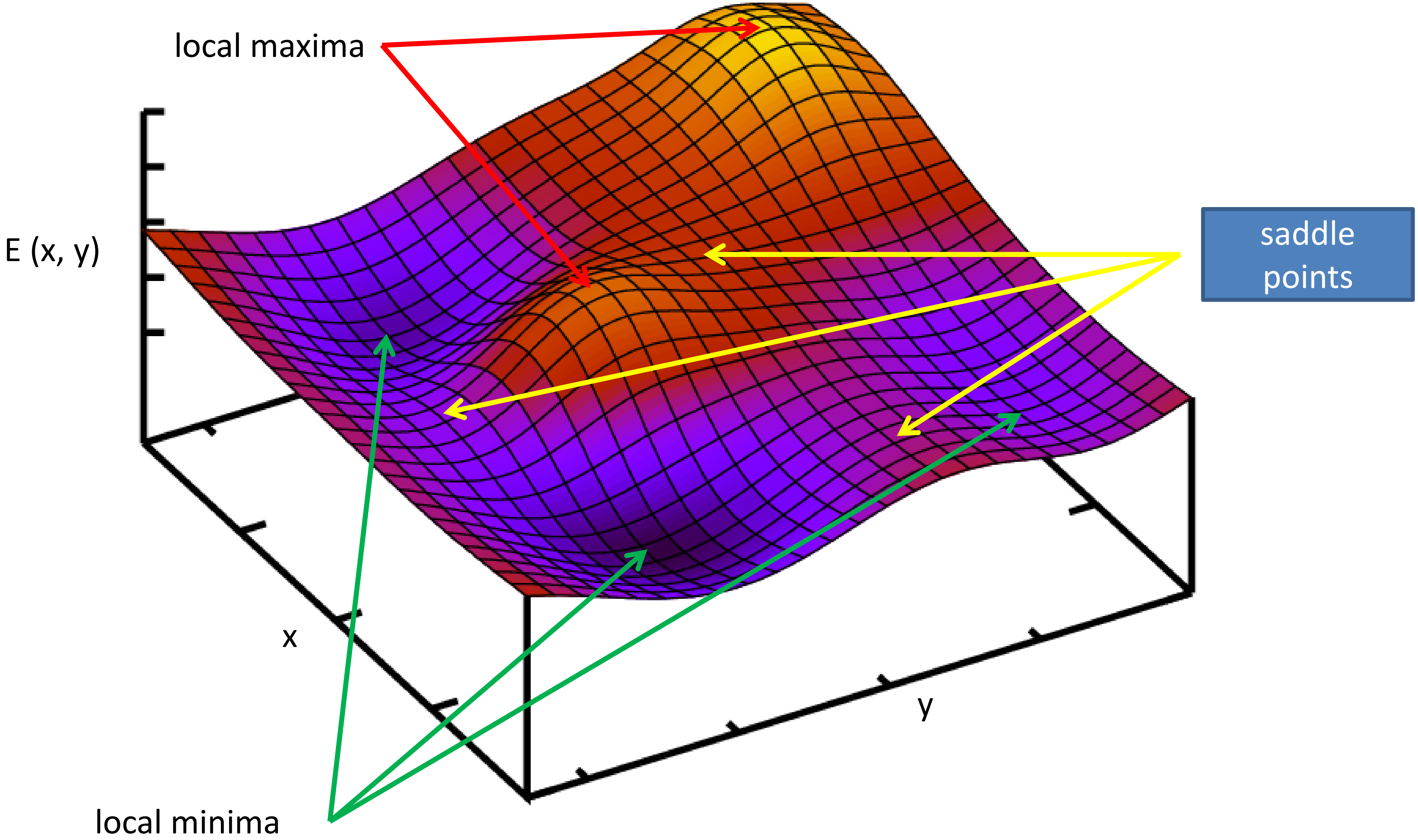
Diagonalization of Hessian is a method for finding eigenvalues and eigenvectors.

- 6 (5) eigenvalues are zero - it corresponds to the translation and rotation of the system
- remaining eigenvalues:
  - **all positive - local minimum**
  - **one negative, other positive - first order saddle point**
  - two negative, other positive - saddle point of the second order
  - .....
  - all negative - local maximum

chemically significant stationary points

N number of ator

Why saddle points and local minima only?



Diagonalization of Hessian

$$\begin{pmatrix} \frac{\partial^2 E(R)}{\partial x_1^2} & \frac{\partial^2 E(R)}{\partial x_1 \partial y_1} & \frac{\partial^2 E(R)}{\partial x_1 \partial z_1} & \cdots & \frac{\partial^2 E(R)}{\partial x_1 \partial z_N} \\ \frac{\partial^2 E(R)}{\partial y_1 \partial x_1} & \frac{\partial^2 E(R)}{\partial y_1^2} & \frac{\partial^2 E(R)}{\partial y_1 \partial z_1} & \cdots & \vdots \\ \frac{\partial^2 E(R)}{\partial z_1 \partial x_1} & \frac{\partial^2 E(R)}{\partial z_1 \partial y_1} & \frac{\partial^2 E(R)}{\partial z_1^2} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E(R)}{\partial z_N \partial x_1} & \cdots & \cdots & \cdots & \frac{\partial^2 E(R)}{\partial z_N^2} \end{pmatrix} = \mathbf{H}(\mathbf{R})$$

Diagonalization of Hessian is an operation, where a **rotation of the coordinate system** is searched such, that the mixed second energy derivatives are zero.
Only the **diagonal elements of the matrix** can be **non-zero** .

Eigenvalues of Hessian then determine the curvature of the function in **direction of axes of the new coordinate system**. These axes are determined by eigenvectors, which are **orthonormal**.

$$\mathbf{c}_i . \mathbf{c}_j = \delta_{ij} \qquad \qquad |\mathbf{c}_k| = 1$$

orthogonal normalization
condition

eigenvalues

$$\begin{pmatrix} \frac{\partial^2 E(R)}{\partial c_1^2} & 0 & 0 & \cdots & 0 \\ 0 & \frac{\partial^2 E(R)}{\partial c_2^2} & 0 & \cdots & \vdots \\ 0 & 0 & \frac{\partial^2 E(R)}{\partial c_3^2} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \frac{\partial^2 E(R)}{\partial c_N^2} \end{pmatrix} = \boldsymbol{\lambda}(\mathbf{R})$$

Energy, Gradient, Hessian

Potential energy calculation

The calculation of the potential energy E(R) is possible by:

- approximate solution of the Schrödinger equation (quantum mechanics, QM)
    - HF method
    - post HF methods (MPn, CI, CC)
    - DFT methods (various functionals)
- using empirical force fields (molecular mechanics, MM)
    - forms and parameters of force fields
- hybrid QM/MM approach
    - interface, type of QM-MM interaction, link atoms, ...
- using coarse grained models

hundreds of methods differing in the approximations used

Energy gradient calculation

Energy gradient:

$$\frac{\partial E(\mathbf{R})}{\partial \mathbf{R}} \qquad \nabla E(\mathbf{R}) \qquad \nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial y_1}, \frac{\partial}{\partial z_1}, \dots, \frac{\partial}{\partial z_N} \right)$$

it is a vector; the number of components is 3N

$$\left(\frac{\partial E(\mathbf{R})}{\partial x_1}, \frac{\partial E(\mathbf{R})}{\partial y_1}, \frac{\partial E(\mathbf{R})}{\partial z_1}, \dots, \frac{\partial E(\mathbf{R})}{\partial z_N} \right)$$

The gradient calculation can be performed:

- **analytically**
- **numerically**

Analytical/Numerical gradient

Analytical gradient calculation is the **preferred method of calculation** in cases where the expression and subsequent calculation of energy derivatives are easy.

computationally more demanding than energy calculation

Numerical gradient calculation is used when the analytical gradient is not available, for example due to the complexity of the implementation of the algorithm for its calculation.

Either **forward differences** (FD) or **central differences** (CD) method can be used to calculate the numerical gradient . In rare cases, it is also possible to use multipoint methods.

The CD method is more accurate than the FD method and therefore the preferred method of gradient calculation.

*FD requires 2*3*N energy calculations*
computationally more demanding than analytical gradient calculation

Calculation of Hessian

Hessian of energy:

$$\begin{pmatrix} \frac{\partial^2 E(R)}{\partial x_1^2} & \frac{\partial^2 E(R)}{\partial x_1 \partial y_1} & \frac{\partial^2 E(R)}{\partial x_1 \partial z_1} & \cdots & \frac{\partial^2 E(R)}{\partial x_1 \partial z_N} \\ \frac{\partial^2 E(R)}{\partial y_1 \partial x_1} & \frac{\partial^2 E(R)}{\partial y_1^2} & \frac{\partial^2 E(R)}{\partial y_1 \partial z_1} & \cdots & \vdots \\ \frac{\partial^2 E(R)}{\partial z_1 \partial x_1} & \frac{\partial^2 E(R)}{\partial z_1 \partial y_1} & \frac{\partial^2 E(R)}{\partial z_1^2} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E(R)}{\partial z_N \partial x_1} & \cdots & \cdots & \cdots & \frac{\partial^2 E(R)}{\partial z_N^2} \end{pmatrix} = \mathbf{H}(\mathbf{R})$$

it is a matrix; number of components is 3Nx3N

Calculation of Hessian can be implemented:

- **analytically** (**memory and computationally intensive**)
- **numerically** (by the method of central differences)
  - from energies (3 x N x 3 x N x 2 energy calculations)
  - from gradients (3 x N x 2 gradient calculations)

Summary

- Quantum states (thermodynamic microstates) from solution of SE are characterized by stationary points on PES.
- Stationary point has zero gradient.

Energy and its gradient MUST be calculated at the same level of theory (model chemistry)!

- Type of stationary point can be determined from PES curvature at the stationary point (Hessian eigenvalue analysis).

Energy, its gradient, and Hessian MUST be calculated at the same level of theory (model chemistry)!

- The most important stationary points are **local minima** (stable states such as reactants, products, intermediates) and **first order saddle points** (transition states).