# Introduction to Geometry Optimization

Dr. Zaheer Ul-Haq

*Associate Professor*

Dr. Panjwani Center for Molecular Medicine and Drug Research

ICCBS, University of Karachi, Pakistan

# Introduction

In the last two decades a new field in chemistry has opened up; experiments on individual molecules have been performed using a number of different techniques e.g. scanning probe microscopy (SPM), atomic force microscopy (AFM) etc.

These experiments investigate:

elastic properties of polymers

conformational changes

the rupture of covalent bonds

and even the formation of new bonds

Manipulations on single molecules involve external forces. The whole field of mechanical manipulation of molecules by applying an external force is known collectively as mechanochemistry.

A molecule exposed to the stresses caused by external forces changes its structure.

How to predict such structural changes ?

# Finding molecular structures

A common theoretical tool used to determine molecular structure is the geometry optimization procedure. In general two types of molecular structure are needed:

the equilibrium geometry $\mathbf{R}_{eq}$ and the transition state geometry $\mathbf{R}_{ts}$

both correspond to the stationary points on the potential energy surface (PES) (molecular energy E($\mathbf{R}$) as a function of nuclear positions R=(R$_1$, R$_2$,…) )

The equilibrium geometry – local minimum
The transition state – saddle point (I-st order)

These points are determined by the condition that the first derivatives of the energy with respect to the nuclei positions vanish (the total force acting on each nucleus vanishes) and the second derivatives are all positive at local minimums and one is negative at saddle points.

# Finding molecular structures

It is propose to use the geometry optimization procedure also to determine enforced structural changes in a molecule.

- In the field of **computational chemistry**, **energy minimization** is the process of finding an arrangement in space of a collection of atoms where, the net inter-atomic force on each atom is acceptably close to zero and the position on the **potential energy surface** (PES) is a stationary point.

- The collection of atoms might be a single **molecule**, an ion, a **condensed phase**, a **transition state** or even a collection of any of these.

- As an example, when optimizing the geometry of a **water molecule**, one aims to obtain the H-H bond lengths and the H-OH bond angle which minimize the forces.

- The motivation for performing a geometry optimization is the physical significance of the obtained structure: optimized structures often correspond to a substance as it is found in nature and the geometry of such a structure can be used in a variety of experimental and theoretical investigations.

- Typically, but not always, the process seeks to find the geometry of a particular arrangement of the atoms that represents a local or global energy minimum.

- Instead of searching for global energy minimum, it might be desirable to optimize to a **transition state**, that is, a saddle point on the potential energy surface. Additionally, certain coordinates (such as a chemical bond length) might be fixed during the optimization.

# Geometry Optimization
## Le Chatliers' Principle

The optimum geometry is the geometry which minimizes the strain on a given system.  Any perturbation from this geometry will induce the system to change so as to reduce this perturbation unless prevented by external forces
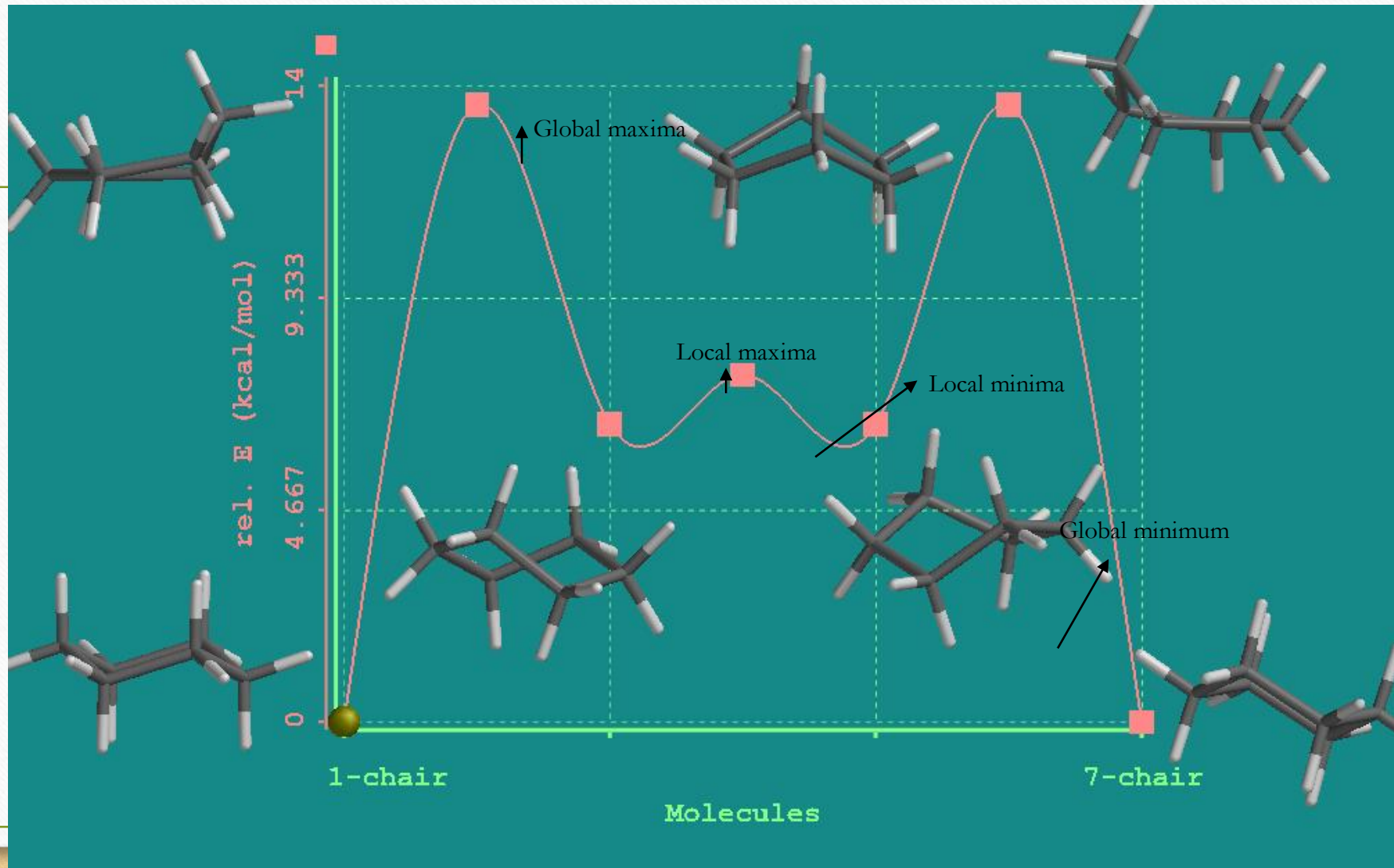
Mathematical Surface Reflects This Principle!!
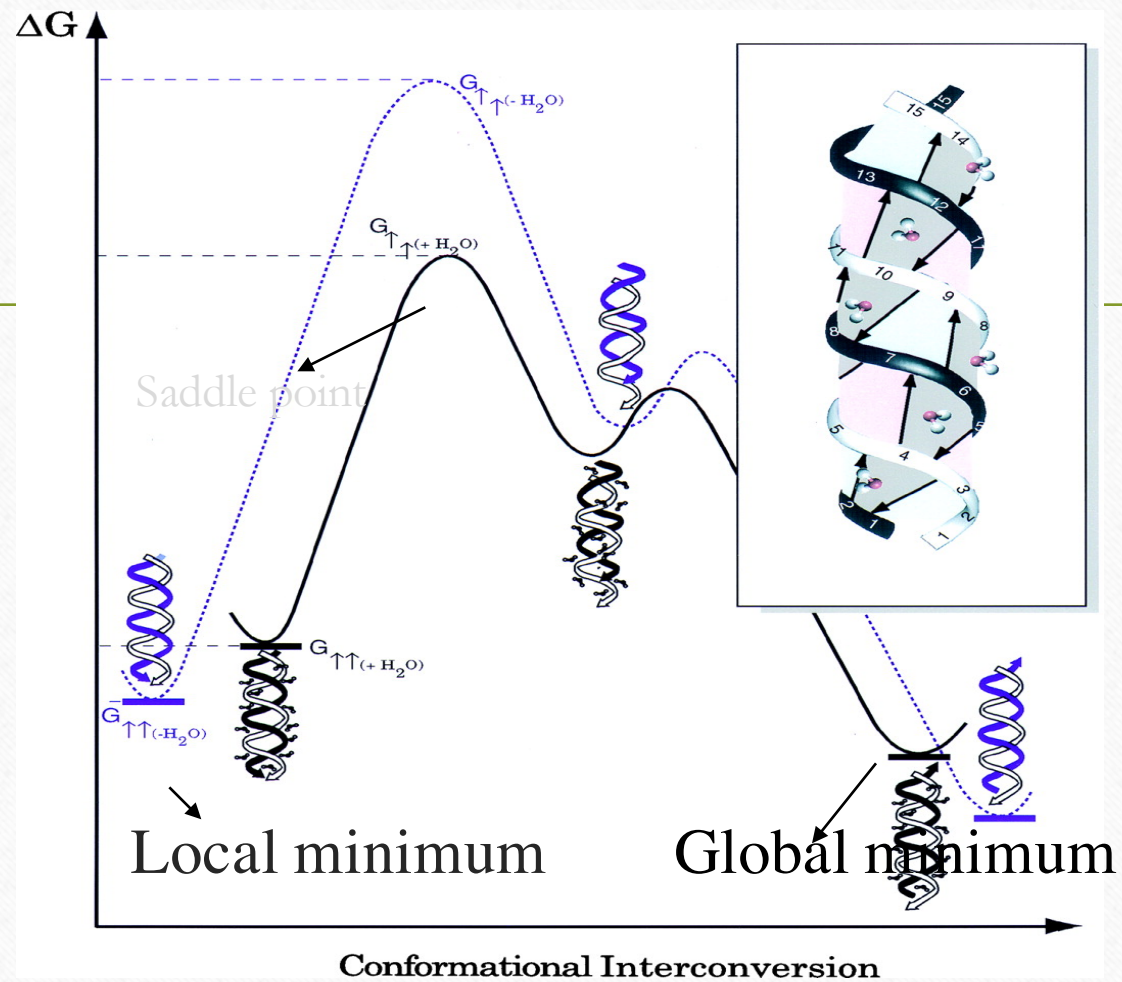
# Features of Potential Energy Surfaces

# Cyclohexane

Local minimum    Global minimum

Conformational Interconversion
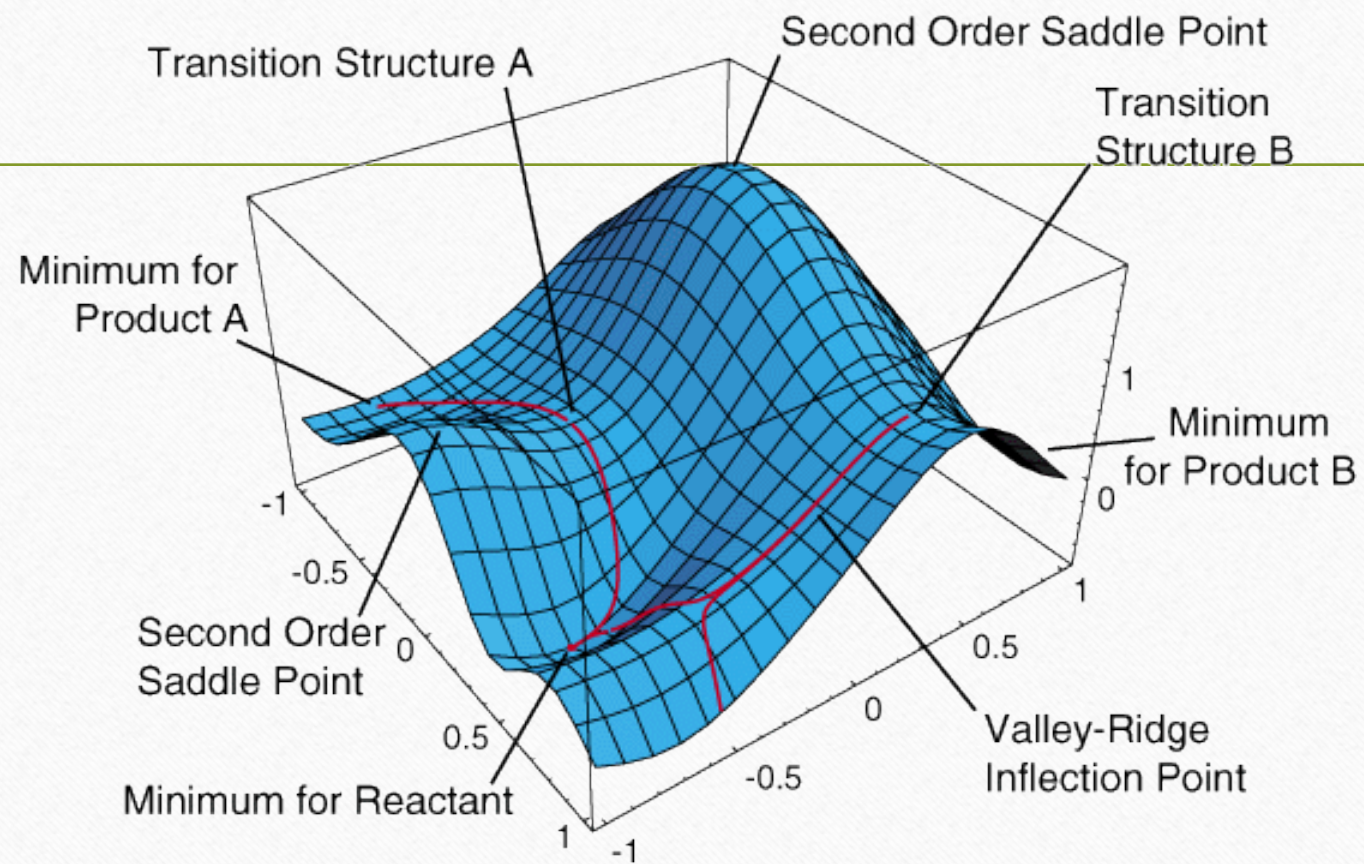
# Potential Energy Surface Terms

- **Gradient** - the first derivative of the energy with respect to geometry (X, Y & Z); also termed the **Force** (strictly speaking, the <u>negative</u> of the gradient is the force)

- **Stationary Points** - points on the PES where the gradient (or force) is zero; this includes **Maxima, Minima, Transition States** (which are first order saddle points), and higher order **Saddle Points**.

# PES Terms…

- In order to distinguish among the latter, one must examine the <u>second derivatives</u> of the PES with respect to geometry; the matrix of these is termed a **Hessian** (or <u>force</u>) matrix.

- Diagonalization of this matrix yields **Eigenvectors** which are <u>normal modes of vibration</u>; the **Eigenvalues** are proportional to the square of the <u>vibrational frequency</u>. (IR spectra can be derived from these)

# Sign of 2nd Derivatives

- The <u>sign</u> of the second derivative can be used to distinguish between **Maxima** and **Minima** on the PES

- **Minima** on the PES have <u>only positive eigenvalues (vibrational frequencies)</u>

- **Maxima** or **Saddle Points** (maximum in one direction but minimum in other directions) have <u>one or more negative</u> (*imaginary*) <u>frequencies</u>.

- A **frequency calculation** must be performed to determine the sign of the vibrational frequencies.

# Molecular geometry and mathematical interpretation

- The geometry of a set of atoms or molecules can be described by Cartesian coordinates of the atoms or, *internal coordinates* formed from a set of bond lengths, bond angles and dihedral angles.

- Given a set of atoms and a vector, **r**, describing the atoms' positions, one can introduce the concept of the energy as a function of the positions, $E(\mathbf{r})$.

- Geometry optimization is then a **mathematical optimization** problem, in which it is desired to find the value of **r** for which $E(\mathbf{r})$ is at a local minimum, that is, the derivative of the energy with respect to the position of the atoms, $\partial E/\partial \mathbf{r}$, is the zero vector.

A special case of a geometry optimization is a search for the geometry of a transition state, and this will be discussed later.

The computational model that provides an approximate $E(\mathbf{r})$ could be based on quantum mechanics (using either density functional theory or semi-empirical methods), force fields, or a combination of those in case of QM/MM.

# Practical aspects of optimization

- Some method such as quantum mechanics can be used to calculate the energy, $E(\mathbf{r})$ , the gradient of the PES, that is, the derivative of the energy with respect to the position of the atoms, $\partial E/\partial \mathbf{r}$ .

- An optimization algorithm can use some or all of $E(\mathbf{r})$ , $\partial E/\partial \mathbf{r}$ and $\partial\partial E/\partial r_i \partial r_j$ to try to minimize the forces and this could in theory be any method such as gradient descent, conjugate gradient or Newton's method.

- For most systems of practical interest, however, it may be prohibitively expensive to compute the second derivative matrix.

The choice of the coordinate system can be crucial for performing a successful optimization.

Cartesian coordinates, for example, are redundant since a non-linear molecule with $N$ atoms has $3N–6$ vibrational degrees of freedom.

Additionally, Cartesian coordinates are highly correlated, that is, the Hessian matrix has many non-diagonal terms that are not close to zero. This can lead to numerical problems in the optimization, because, for example, it is difficult to obtain a good approximation to the Hessian matrix and calculating it precisely is too computationally expensive.

Internal coordinates tend to be less correlated but are more difficult to set-up and it can be difficult to describe some systems, such as ones with symmetry or large condensed phases.

# Degree of freedom restriction

- Some degrees of freedom can be eliminated from an optimization, for example, positions of atoms or bond lengths and angles can be given fixed values. Sometimes these are referred to as being *frozen* degrees of freedom.

# Transition state optimization

- Transition state structures can be determined by searching for saddle points on the PES.

- A first-order saddle point is a position on the PES corresponding to a minimum in all directions except one; a second-order saddle point is a minimum in all directions except two, and so on.

- Algorithms to locate transition state geometries fall into two main categories: local methods and semi-global methods.

- Local methods are suitable when the starting point for the optimization is very close to the true transition state and semi-global methods find application when it is sought to locate the transition state with very little *a priori* knowledge of its geometry. Some methods, such as the Dimer method, fall into both categories.

# Local searches

- A so-called local optimization requires an initial guess of the transition state. Initial guess must have a corresponding Hessian matrix with one negative Eigenvalue, or, the negative Eigenvalue corresponding to the reaction coordinate must be greater in magnitude than the other negative Eigenvalues.

- Given the above pre-requisites, a local optimization algorithm can then move "uphill" along the Eigenvector with the most negative Eigenvalue and "downhill" along all other degrees of freedom, using something similar to a quasi-Newton method.

# Dimer method

- The dimer method can be used to find possible transition states without knowledge of the final structure or to refine a good guess of a transition structure. The "dimer" is formed by two images very close to each other on the PES. The method works by moving the dimer uphill from the starting position whilst rotating the dimer to find the direction of lowest curvature (ultimately negative).

# Activation Relaxation Technique (ART)

- The Activation Relaxation Technique (ART) is also an open-ended method to find new transition states or to refine known saddle points on the PES. The method follows the direction of lowest negative curvature (computed using the Lanczos algorithm) on the PES to reach the saddle point, relaxing in the perpendicular hyperplane between each "jump" (activation) in this direction.

# Chain-of-state methods

- Chain-of-state methods can be used to find the *approximate* geometry of the transition state based on the geometries of the reactant and product. The generated approximate geometry can then serve as a starting point for refinement via a local search.

- Chain-of-state methods use a series of vectors, that is points on the PES, connecting the reactant and product of the reaction of interest, $\mathbf{r}_{reactant}$ and $\mathbf{r}_{product}$, thus discretizing the reaction pathway. Very commonly, these points are referred to as *beads* due to an analogy of a set of beads connected by strings or springs, which connect the reactant and products. The series of beads is often initially created by interpolating between $\mathbf{r}_{reactant}$ and $\mathbf{r}_{product}$,

- for example, for a series of $N + 1$ beads, bead $i$ might be given by

$$\mathbf{r}_i = \frac{i}{N}\mathbf{r}_{\text{product}} + \left(1 - \frac{i}{N}\right)\mathbf{r}_{\text{reactant}}$$

- where $i \in 0, 1, ..., N$. Each of the beads $\mathbf{r}_i$ has an energy, $E(\mathbf{r}_i)$, and forces, $-\partial E/\partial \mathbf{r}_i$ and these are treated with a constrained optimization process that seeks to get an as accurate as possible representation of the reaction pathway. For this to be achieved, spacing constraints must be applied so that each bead $\mathbf{r}_i$ does not simply get optimized to the reactant and product geometry.

- Often this constraint is achieved by projecting out components of the force on each bead $\mathbf{r}_i$, or alternatively the movement of each bead during optimization, that are tangential to the reaction path. For example, if for convenience, it is defined that $\mathbf{g}_i = \partial E/\partial \mathbf{r}_i$, then the energy gradient at each bead minus the component of the energy gradient that is tangential to the reaction pathway is given by

$$\mathbf{g}_i^{\perp} = \mathbf{g}_i - \tau_i(\tau_i \cdot \mathbf{g}_i) = \left(I - \tau_i\tau_i^{T}\right)\mathbf{g}_i$$

- where *I* is the identity matrix and $\boldsymbol{\tau}_i$ is a unit vector representing the reaction path tangent at $\mathbf{r}_i$. By projecting out components of the energy gradient or the optimization step that are parallel to the reaction path, an optimization algorithm significantly reduces the tendency of each of the beads to be optimized directly to a minimum.

# Synchronous transit

- The simplest chain-of-state method is the linear synchronous transit (LST) method. It operates by taking interpolated points between the reactant and product geometries and choosing the one with the highest energy for subsequent refinement via a local search. The quadratic synchronous transit (QST) method extends LST by allowing a parabolic reaction path, with optimization of the highest energy point orthogonally to the parabola.

# Nudged elastic band

In Nudged elastic band method, the beads along the reaction pathway have simulated spring forces in addition to the chemical forces, $-\partial E/\partial \mathbf{r}_i$, to cause the optimizer to maintain the spacing constraint. Specifically, the force $\mathbf{f}_i$ on each point $i$ is given by

$$\mathbf{f}_i = \mathbf{f}_i^{\parallel} - \mathbf{g}_i^{\perp}$$

where

$$\mathbf{f}_i^{\parallel} = k\left[\left((\mathbf{r}_{i+1} - \mathbf{r}_i) - (\mathbf{r}_i - \mathbf{r}_{i-1})\right) \cdot \tau_i\right]\tau_i$$

is the spring force parallel to the pathway at each point $\mathbf{r}_i$ ($k$ is a spring constant and $\tau_i$, is a unit vector representing the reaction path tangent at $\mathbf{r}_i$).

In a traditional implementation, the point with the highest energy is used for subsequent refinement in a local search. There are many variations on the NEB (nudged elastic band) method, such including the climbing image NEB, in which the point with the highest energy is pushed upwards during the optimization procedure so as to give a geometry which is even closer to that of the transition state.

# String method

- The string method uses splines connecting the points, $\mathbf{r}_i$, to measure and enforce distance constraints between the points and to calculate the tangent at each point. In each step of an optimization procedure, the points might be moved according to the force acting on them perpendicular to the path, and then, if the equidistance constraint between the points is no-longer satisfied, the points can be redistributed, using the spline representation of the path to generate new vectors with the required spacing.

- Variations on the string method include the growing string method, in which the guess of the pathway is grown in from the end points (that is the reactant and products) as the optimization progresses.

# Comparison with other techniques

- Geometry optimization is fundamentally different from a molecular dynamics simulation. The latter simulates the motion of molecules with respect to time, subject to temperature, chemical forces, initial velocities, Brownian motion of a solvent, and so on, via the application of Newton's laws of Motion. This means that the trajectories of the atoms which get computed, have some physical meaning. Geometry optimization, by contrast, does not produced a "trajectory" with any physical meaning – it is concerned with minimization of the forces acting on each atom in a collection of atoms, and the pathway via which it achieves this is lacks meaning. Different optimization algorithms could give the same result for the minimum energy structure, but arrive at it via a different pathway.

# OPTIMIZATION METHODS

# Methods of Optimisation

- Energy only:
  - simplex
- Energy and first derivatives (forces):
  - steepest descents (poor convergence)
  - conjugate gradients (retains information)
  - approximate Hessian update
- Energy, first and second derivatives
  - Newton-Raphson
  - Broyden (BFGS) updating of Hessian (reduces inversions)
  - Rational Function Optimisation (for transition states/ and soft modes)

# Energy Only (Univariate) Method

- Simplest to implement

- Proceeds one direction until energy increases, then turns 90°, etc.

- Least efficient
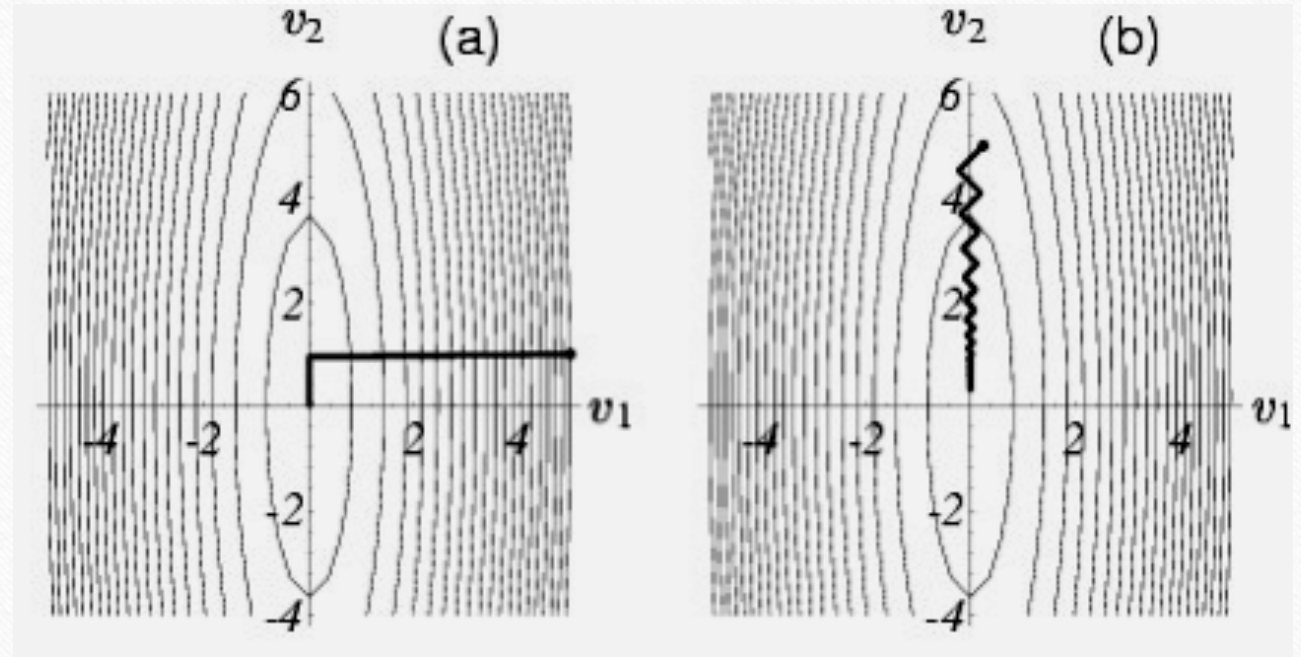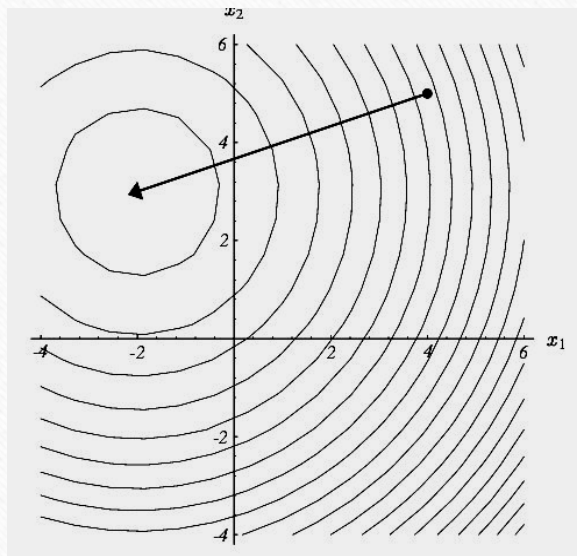  - many steps
  - steps are not guided

- Not used very much.



Univariate Search
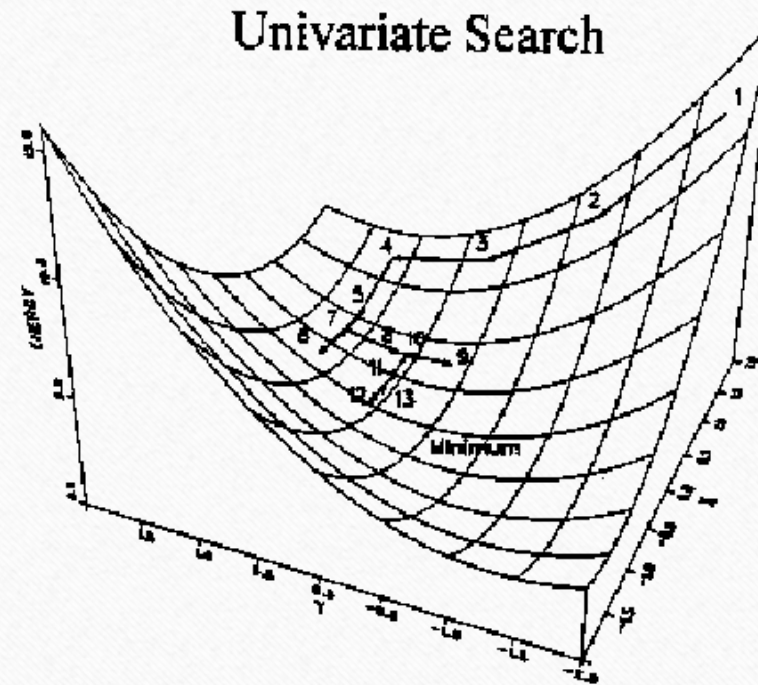
# Steepest Descents

- **Simplest method in use**
- **Start at $x_o$**
- **Minimize f(x) along the line defined by the gradient**
- **Follows most negative gradient (max. force)**
- **Fastest method from a poor starting geometry**
- **Converges slowly near the energy minimum, start again until tolerance is reached**
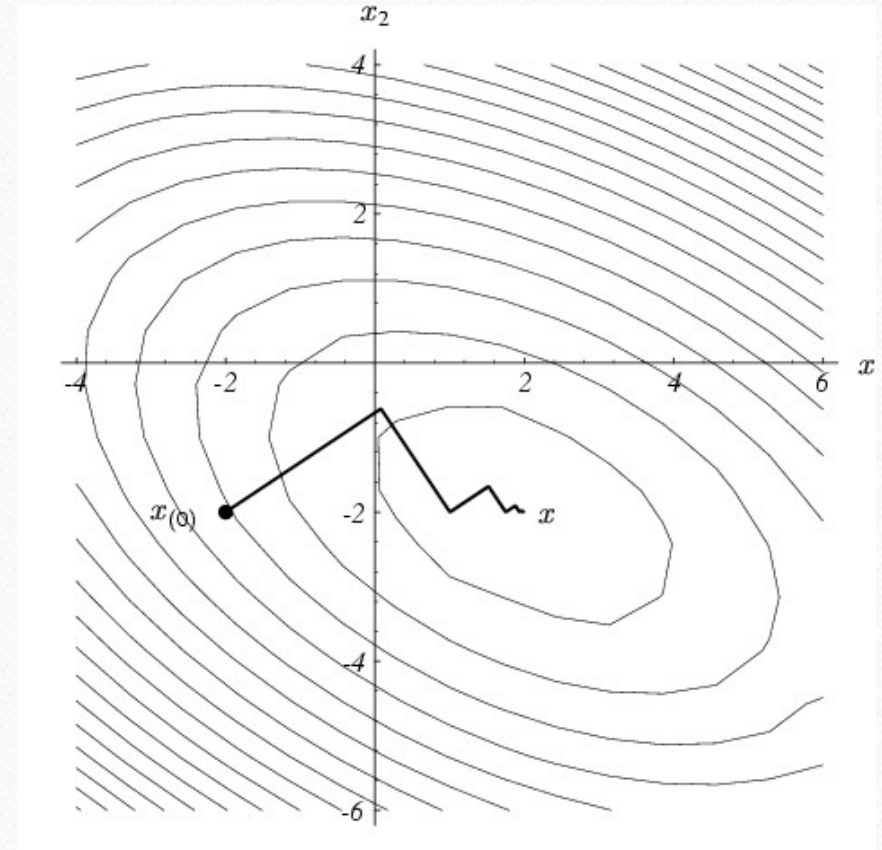
# Steepest Descents



**Performance depends on**
- **Eigenvalues of Hessian ($\lambda_{max}$ / $\lambda_{min}$)**
- **Starting point**

# Conjugate Gradients

- Same idea, but retaining information about previous steps

- Search directions 'conjugate' (orthogonal) to previous

- Convergence assured for N steps

- Variations on this procedure are the **Fletcher-Reeves**, the **Davidon- Fletcher-Powell** and the **Polak-Ribiere** methods.

# Second Derivative Methods

- The 2nd derivative of the energy with respect to X,Y,Z [the Hessian] determines the pathway.

- Computationally more involved, but generally fast and reliable, esp. near the minimum.



Quasi-Newton
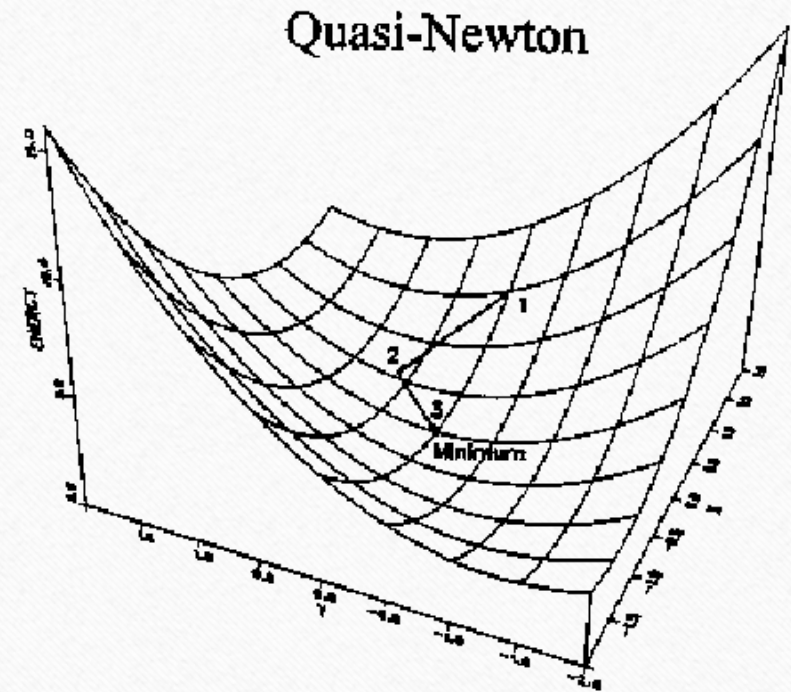
# Geometry Optimization (Summary)

- Optimum structure gives useful information

$$\frac{dV(\vec{r})}{d\vec{r}} = 0$$

- First Derivative is Zero - At minimum/**maximum**

$$\frac{d^2V(\vec{r})}{d\vec{r}^2} > 0, \min$$

- **Use Second Derivative to establish minimum/maximum**

$$\frac{d^2V(\vec{r})}{d\vec{r}^2} < 0, \max$$

- As N increases so does dimensionality/complexity/beauty/difficulty

# Geometry Optimization (Summary)

- Method used depends on
  - System size
    - 1-d (line search, bracketing, steepest descent)
    - N-d local (Downhill)
      - W/o derivatives
        - Simplex
        - Direction set methods (Powell's)
      - W/ derivatives
        - Conjugate gradient
        - Newton or variable metric methods
    - N-d Global
      - Monte Carlo
      - Simulated Annealing
      - Genetic Algoritms
  - Form of energy
    - Analytic
    - Not analytic