



# Molecular Mechanics

A Brief Introduction



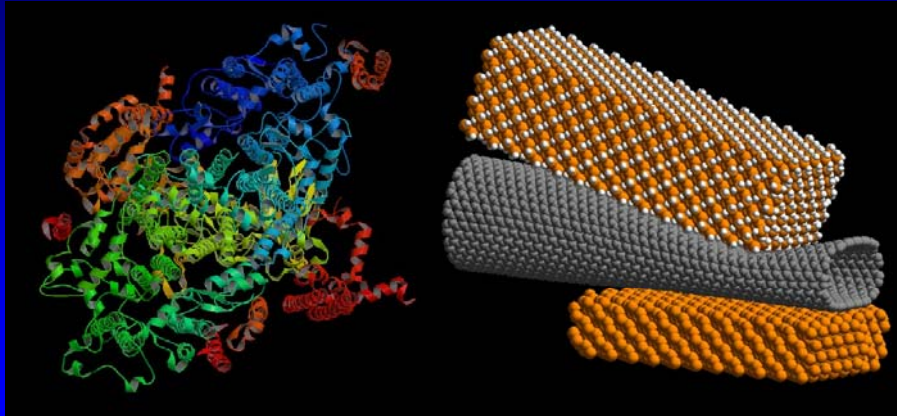
## The challenge: Large systems

➤ Last year: the modeling of electronic properties from first-principles approach (HF)

The problem: computer resource ( $t \propto N^3$ )

Overshooting some problems  
(Too many details)

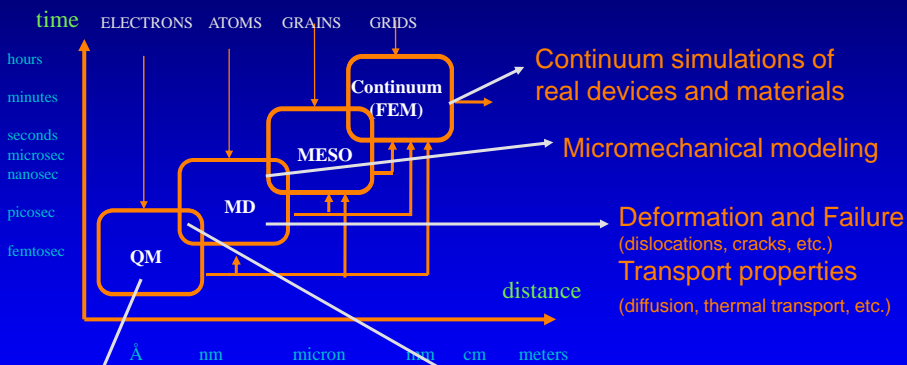
# Modeling....



Biology & Biochemistry

Material Science

# Scale factor



Accurate calculations for bulk phases and molecules (EOS, dissociation curves, etc.)

Optical properties

New generation reactive force fields based purely on first principles

For metals, oxides, organics.  
Describes: mechanical properties, chemistry, charge transfer, atomic polarization, etc.



## Why Molecular Mechanics

The "mechanical" molecular model was developed out of a need to describe molecular structures and properties in as practical a manner as possible. In particular, it has been developed for large systems, even if application to « small » organic molecule is possible



## Molecular Mechanics

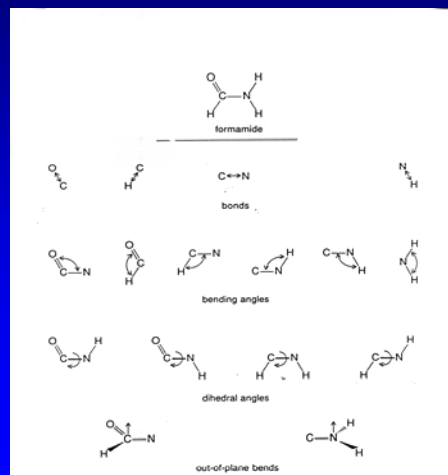
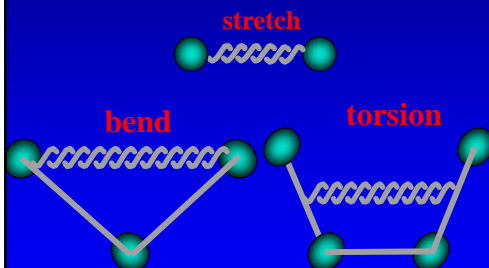
- Calculation of the lowest energy molecular structure, based on principles of classical (Newtonian) physics
- Based on additive energy contributions due to deviation from some "ideal" geometry
- Contributions include based on chemical ideas (bond stretching, bond angle bending, torsional angle deformation, dipole-dipole interactions, van der Waals forces, H-bonding and other terms).

## Main assumptions

- Nuclei and electrons are modelled as atom-like particles.
- Atom-like particles are spherical (radii obtained from measurements or theory) and have a net charge (obtained from theory).
- Interactions are additive
- Interactions are based on springs and classical potentials.
- Interactions must be preassigned to specific sets of atoms.
- Interactions determine the **spatial distribution** of atom-like particles and their **energies**.

## Molecular decomposition

A Molecule is a Collection of Atoms Held Together by Forces



Forces are Described by Potential Energy Functions

## Bond Stretching Potential

- A Morse potential best describes energy of bond stretching (& compression), but it is too complex for efficient calculation and it requires three parameters for each bond.

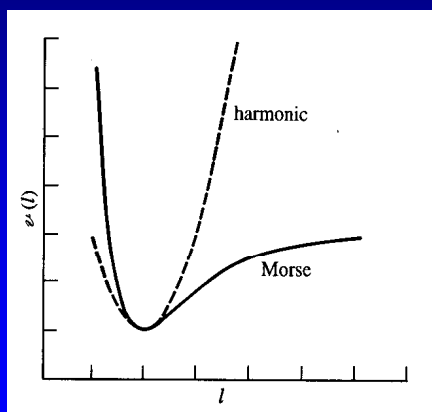
$$v(l) = D_e \{1 - \exp[-a(l - l_0)]\}^2$$

if:  $D_e$  = depth of potential energy minimum,

$a = \omega\sqrt{\mu/2D_e}$  where  $\mu$  is the reduced mass and  $\omega$  is related to the bond stretching frequency by

$$\omega = \sqrt{k/\mu}$$

## Morse potential & Hooke's Law



- Most bonds deviate in length very little from their equilibrium values, so simpler mathematical expressions, such as the harmonic oscillator (Hooke's Law) have been used to model the bond stretching energy:

$$v(l) = k/2(l - l_0)^2$$



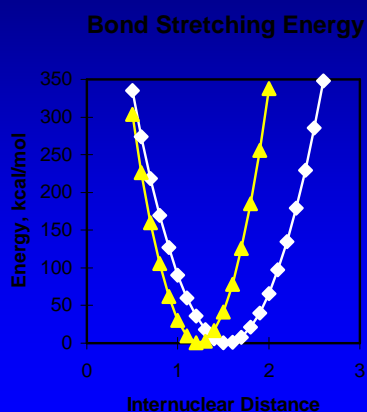
## Bond Stretching Potential

$$E_{\text{stretch}} = k_s/2 (l - l_0)^2$$

(Hooke's law force...

harmonic oscillator)

graph: C-C; C=O



## Bond Stretching Parameters (MM2)

Bond	$l_0$ (Å)	$k$ (kcal mol <sup>-1</sup> Å <sup>-2</sup> )
Csp <sup>3</sup> -Csp <sup>3</sup>	1.523	317
Csp <sup>3</sup> -Csp <sup>2</sup>	1.497	317
Csp <sup>2</sup> =Csp <sup>2</sup>	1.337	690
Csp <sup>2</sup> =O	1.208	777
Csp <sup>3</sup> -Nsp <sup>3</sup>	1.438	367
C-N (amide)	1.345	719

- Hard mode.
- Bond types correlate with  $l_0$  and  $k$  values.
- A 0.2Å deviation from  $l_0$  when  $k=300$  leads to an energy increase of 12 kcal/mol.

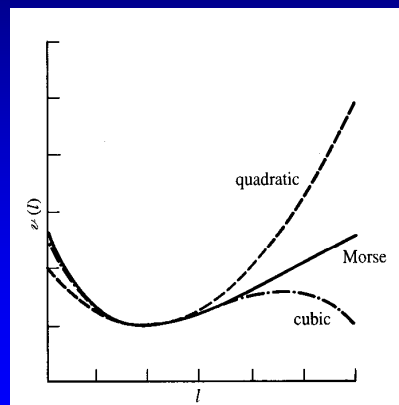


## Higher order terms give better fit

With cubic and higher terms:

$$v(l) = k/2(l - l_0)^2 [1 - k'(l - l_0) - k''(l - l_0)^2 - k'''(l - l_0)^3 - \dots]$$

(cubic terms give better fit in region near minimum; inclusion of a fourth power term eliminates the maximum in the cubic fcn.)



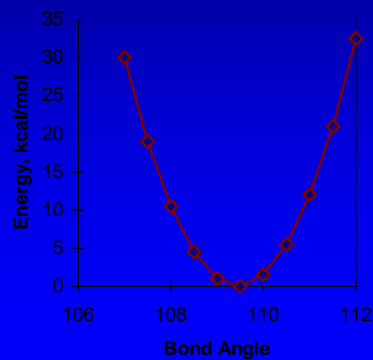
## Bond Angle Bending Potential

$$E_{\text{bend}} = k_b/2 (\theta - \theta_0)^2$$

graph:  $sp^3$  C-C-C

(Likewise, cubic and higher terms are added for better fit with experimental observations)

Bond Angle Deformation, C-C-C



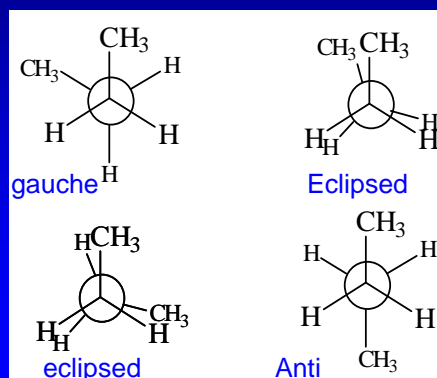
## Angle Bending Parameters (MM2)

Angle	$\theta_0$	$k$ (kcal mol <sup>-1</sup> deg <sup>-1</sup> )
Csp <sup>3</sup> -Csp <sup>3</sup> -Csp <sup>3</sup>	109.47	0.0099
Csp <sup>3</sup> -Csp <sup>3</sup> -H	109.47	0.0079
H-Csp <sup>3</sup> -H	109.47	0.007
Csp <sup>3</sup> -Csp <sup>2</sup> -Csp <sup>3</sup>	117.2	0.0099
Csp <sup>3</sup> -Csp <sup>2</sup> =Csp <sup>2</sup>	121.4	0.0121
Csp <sup>3</sup> -Csp <sup>2</sup> =O	122.5	0.0101

- Hard mode (softer than bond stretching)

## Torsional Potential

- Related to the rotation “barrier” (which also includes some other contributions, such as van der Waals interactions).
- The potential energy increases periodically as eclipsing interactions occur during bond rotation.

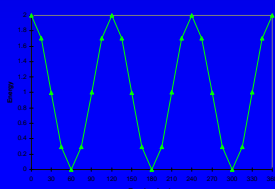
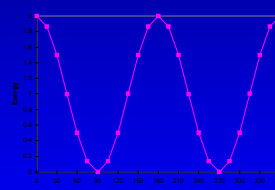
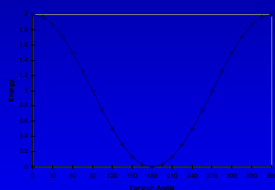




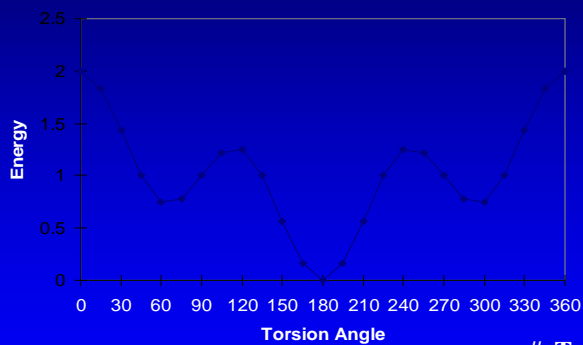


# Torsional Potential

$$E_{\text{torsion}} = 0.5 V_1 (1 + \cos \phi) + 0.5 V_2 (1 + \cos 2\phi) + 0.5 V_3 (1 + \cos 3\phi)$$



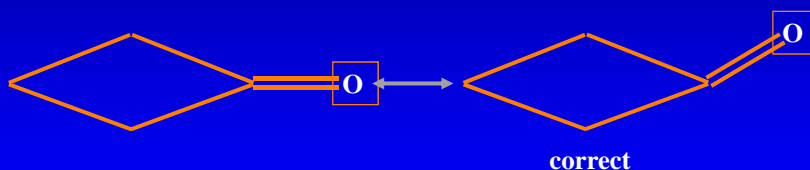
# Butane Barrier



#	Type	V1	V2	V3
1	C-C-C-C	0.200	0.270	0.093
4	C-C-C-H	0.000	0.000	0.267
4	H-C-C-H	0.000	0.000	0.237

## Out-of-Plane Bending

- Allows for non-planarity when required (*e.g.*, cyclobutanone).
- Prevents inversion about chiral centers (*e.g.*, for united atoms).
- Induces planarity when required.



## Out-of-Plane Bending

### ■ Wilson Angle

Between *ijk* plane and *i-l* bond

$$v(\chi) = k(\chi - \chi_0)^2$$

### ■ Pyramidal Distance

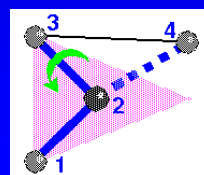
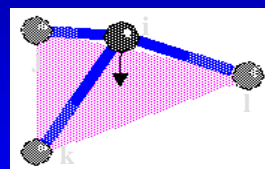
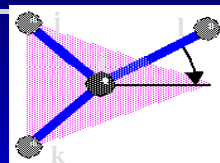
Between atom *i* and *jk* plane

$$v(d) = k(d - d_0)^2$$

### ■ Improper torsion

$$v(\chi) = k[1 + \cos(n\chi - \chi_0)]$$

1-2-3-4 torsion

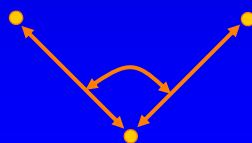


## Cross Terms: Stretch-Bend

- Coupling between internal coordinates.
- Important for reproducing structures of unusual (*e.g.*, highly strained ) systems and of vibrational spectra.

**Stretch-Bend:** As a bond angle decreases, the adjacent bonds stretch to reduce the interaction between the 1,3 atoms

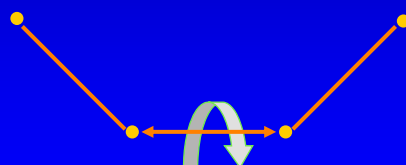
$$v(l_1, l_2, \theta) = \frac{k_{l_1 l_2}}{2} [(l_1 - l_{1,0}) + (l_2 - l_{2,0})](\theta - \theta_0)$$



## Cross Terms: Stretch-Torsion

**Stretch-Torsion:** For an A-B-C-D torsion, the central B-C bond elongates in response to eclipsing of the A-B and C-D terminal bonds.

$$v(l, \omega) = k(l - l_0) \cos n\omega$$



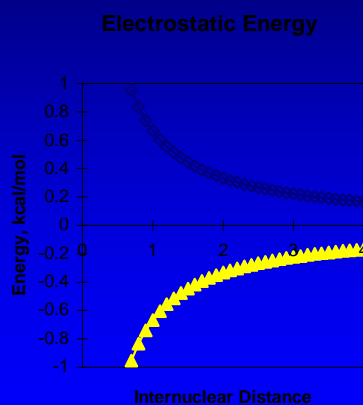
## Non-Bonded Interactions

- Operate within molecules and between molecules.
- Through space interactions.
- Modeled as a function of an inverse power of the distance.
- Soft mode.
- Divided into:
  - Electrostatic interactions.
  - VdW interactions.

## Electrostatic Interaction

$$E_{\text{electrostatic}} = q_1 q_2 / c \epsilon r$$

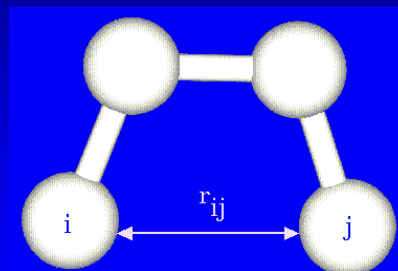
(attractive or repulsive, depending on relative signs of charge; value depends inversely on **permittivity of free space**, or the dielectric constant of the hypothetical medium)



# Electrostatic Energy

Classical Coulomb's Law

$$E_{\text{electrostatic}} = q_i q_j / \epsilon r_{ij}$$

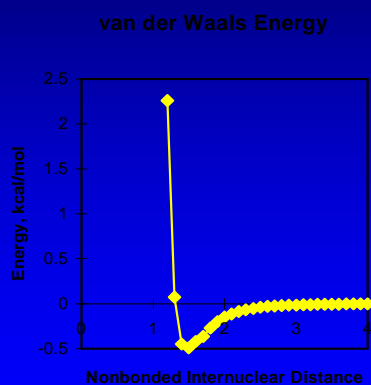


# van der Waals Potential

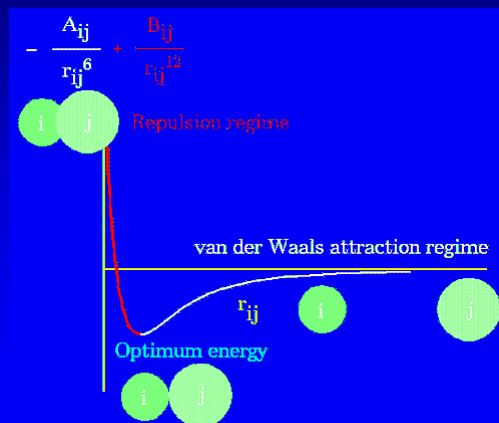
$$E_{\text{vdW}} = A/r^{12} - B/r^6$$

Lennard-Jones or  
6-12 potential

combination of a repulsive  
term [A] and an attractive term [B]



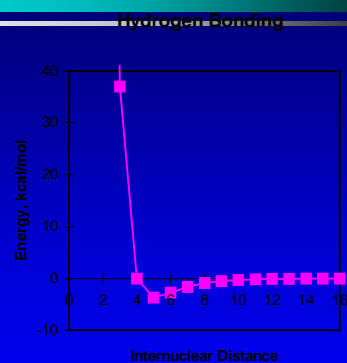
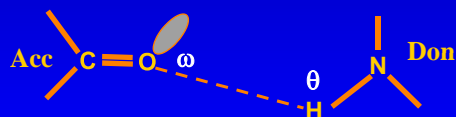
# van der Waals Potential



# Hydrogen Bonding Energy

$$E_{\text{H-Bond}} = A/r^{12} - B/r^{10}$$
 (Lennard-Jones type, 12-10)

- H-bond geometry dependent (co-linearity with lp preferred):



$$v_{HB} = \left( \frac{A}{r_{H...Acc}^{12}} - \frac{B}{r_{H...Acc}^{10}} \right) \cos^2 \theta_{Don...H...Acc} \cos^4 \omega_{H...Acc...LP}$$



## Dipole-Dipole Energy

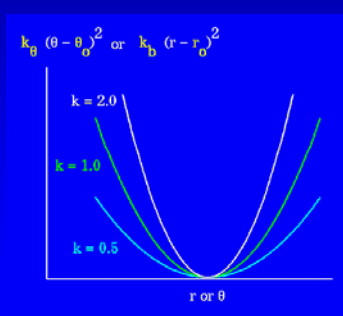
Calculated as the vector sum of the bond dipole moments, also considering the **permittivity** (related to dielectric constant) of the medium (typical default value is 1.5)

(this is too complicated to demonstrate!!!)

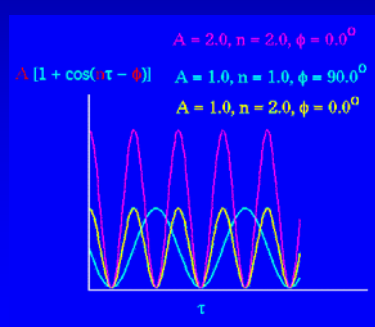


## Parameters effects on potentials

### Stretching



### Torsion



## Force Field

### A Force Field is a Collection of Potential Energy Functions

- A force field is defined by the functional forms of the energy functions and by the values of their parameters.

$$V(r^N) = \sum_{\text{bonds}} \frac{k_i}{2} (l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\ + \sum_{i=1}^N \sum_{j=i+1}^N \left( 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) + \text{cross terms}$$

## MM Energy (Steric Energy)

- A molecular mechanics program will return an energy value for every conformation of the system.
- Steric energy is the energy of the system relative to a reference point. This reference point depends on the bonded interactions and is both force field dependent and molecule dependent.
- Thus, steric energy can only be used to compare the relative stabilities of different conformations of the same molecule and can not be used to compare the relative stabilities of different molecules.
- Further, all conformational energies must be calculated with the same force field.



## Steric Energy

$$\text{Steric Energy} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} + E_{\text{vdW}} + E_{\text{electrostatic}} + E_{\text{stretch-bend}} + E_{\text{torsion-stretch}} + \dots$$

- $E_{\text{stretch}}$  Stretch energy (over all bonds)
- $E_{\text{bend}}$  Bending energy (over all angles)
- $E_{\text{torsion}}$  Torsional (dihedral) energy (over all dihedral angles)
- $E_{\text{vdW}}$  Van Der Waals energy (over all atom pairs > 1,3)
- $E_{\text{electrostatic}}$  Electrostatic energy (over all charged atom pairs >1,3)
- $E_{\text{stretch-bend}}$  Stretch-bend energy
- $E_{\text{torsion-stretch}}$  Torsion-stretch energy

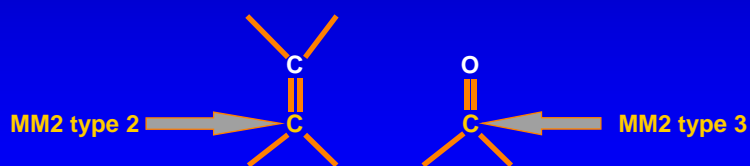
$E_{\text{vdW}} + E_{\text{electrostatic}}$  are often referred to as non-bonded energies

## Force Fields: General Features

- **Force field definition**
  - Functional form (usually a compromise between accuracy and ease of calculation).
  - Parameters (transferability assumed).
- **Force fields are empirical**
  - There is no “correct” form of a force field.
  - Force fields are evaluated based solely on their performance
- **Force field are parameterized for specific properties**
  - Structural properties
  - Energy
  - Vibrational Spectra

## Force Fields: Atom Types

- In molecular mechanics atoms are given types - there are often several types for each element.
- **Atom types depend on:**
  - Atomic number (*e.g.*, C, N, O, H).
  - Hybridization (*e.g.*, SP<sup>3</sup>, SP<sup>2</sup>, SP).
  - Environment (*e.g.*, cyclopropane, cyclobutane).



- “Transferability” is assumed - for example that a C=O bond will behave more or less the same in all molecules.

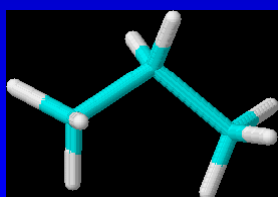
## A General Force Field Calculation

- **Input**
  - Atom Types
  - Starting geometry
  - Connectivity
- **Energy minimization / geometry optimization**
- **Calculate molecular properties at final geometry**
- **Output**
  - Molecular structure
  - Molecular energy
  - Dipole moments
  - etc. etc. etc.

## A Simple MM/FF Calculation

$$V(r^N) = \sum_{\text{bonds}} \frac{k_i}{2} (l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma))$$

$$+ \sum_{i=1}^N \sum_{j=i+1}^N 4\epsilon_{ij} \left\{ \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}$$



### ■ Bonds

- C-C x 2
- C-H x 8

### ■ Angles

- C-C-C x 1
- C-C-H x 10
- H-C-H x 7

### ■ Torsions

- H-C-C-H x 12
- H-C-C-C x 6

### ■ Non-bonded

- H-H x 21
- H-C x 6

## Force Field Parameterization

- Choosing values for the parameters in the potential function equations to best reproduce experimental data.

### ■ Parameterization techniques

- Trial and error
- Least square methods

### ■ Types of parameters

- Stretch: natural bond length ( $l_0$ ) and force constants ( $k$ ).
- Bend: natural bond angles ( $\theta_0$ ) and force constants ( $k$ ).
- Torsions:  $V_i$ 's.
- VdW: ( $\epsilon$ , VdW radii).
- Electrostatic: Partial atomic charges.
- Cross-terms: Cross term parameters.

## MM2 Atom Types (a few of >60)

1	C	sp <sup>3</sup> carbon	13	Br	bromine
2	C	sp <sup>2</sup> carbon (C=C)	14	I	iodine
3	C	sp <sup>2</sup> carbon (C=O)	15	S	sulfide (-S-)
4	C	sp carbon	16	S+	sulfonium
5	H	hydrogen (see others)	17	S	sulfoxide (use S=O)
6	O	oxygen (single bonded)	18	S	sulfone (use two S=O)
7	O	oxygen (double bonded)	19	Si	silane
8	N	sp <sup>3</sup> nitrogen	20	LP	lone pair of electrons
9	N	sp <sup>2</sup> nitrogen	21	H	hydroxyl hydrogen
10	N	sp nitrogen	22	C	cyclopropane carbon
11	F	fluorine	23	H	amine hydrogen
12	Cl	chlorine	24	H	carboxylic acid hydrogen

## Parameters - Quantity

### ■ For N atom types require:

- N non-bonded parameters
- N\*N stretch parameters
- N\*N\*N bend parameters
- N\*N\*N\*N torsion parameters

### ■ Example

MacroModel: 39 atom types

	Required	Actual
Stretch	1521	164
Bend	58,319	357
Torsion	2,313,441	508



## Parameters - Source

- A force field parameterized according to data from one source (*e.g.*, experimental gas phase, experimental solid phase, *ab initio*) will fit data from other sources only qualitatively.
- **Experiment: geometries and non-bonded parameters**
  - X-Ray crystallography
  - Electron diffraction
  - Microwave spectroscopy
  - Lattice energies
- **Advantages**
  - Real
- **Disadvantages**
  - Hard to obtain
  - Non-uniform
  - Limited availability



## Parameters - Source

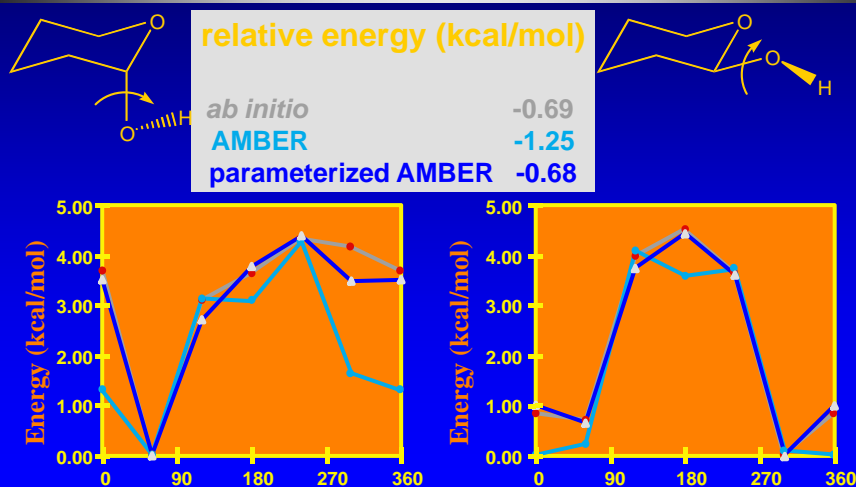
- **High level molecular orbital calculations**
  - Conformational preferences and partial charges (HF/6-31G\*).
  - Corrections for electron correlation effects are often important: MP2, MP3 etc.
  - Charges obtained by fitting electrostatic potentials
- **Advantages**
  - Relatively easy to obtain.
  - Uniform.
  - Unlimited availability.
  - Complete potential energy surfaces are available.
- **Disadvantages**
  - Model dependent.
  - Computationally expensive

## Parameters - Quality

- “Unusual” functional groups or combinations of functionality will probably not be well described by existing force fields.
- The success of modeling with molecular mechanics rests on the use of parameters which are “high quality” *e.g.* not generalized.
- Each standard force field has been designed with a particular (or few) molecular property and target group of compounds in mind.

**Know your parameters!!!!**

## Parameterization - Example



## Transferability of Parameters

- **Assumption: The same set of parameters can be used to model a related series of compounds:**
  - Prediction
  - Missing parameters
    - Educated guess
    - Parameters derived from atomic properties (*e.g.*, UFF)
  - Reducing number of parameters
    - Generalized torsions (*e.g.*, depend only on two central atoms)
    - Generalized VdW parameters (same for SP, SP<sup>2</sup>, SP<sup>3</sup>)
- **Above assumption breaks down for close interaction function groups:**



## Properties Calculated

- ↗ Optimized Structures (minimum energy conformation)
- ↗ Dipole moment (vector sum of bond dipoles)
- ↗ Enthalpy of Formation
- ↗ Vibrational Spectra

## Performance of Molecular Mechanics

- Molecular mechanics calculates the following molecular features to be within the listed numbers from the experimental results.

	MM2	MM3	Dreiding	CFF
Bond length ( $\pm$ A)	0.03	0.01	0.03	0.01
Bond angle ( $\pm$ Deg.)	1	1	3	2
Torsion angle ( $\pm$ Deg.)	8	8	8	
Dipole ( $\pm$ Debye)	0.1	0.07		
$\Delta H_f$ ( $\pm$ kcal/mol)	0.5	0.6		2
Vibrational frequencies ( $\pm$ )	80	40		

## Common Force Fields

- MM2 / MM3 (Allinger) best; general purpose
- MMX (Gilbert) added TS's, other elements; good
- MM+ (Ostlund) in HyperChem; general; good
- OPLS (Jorgenson) proteins and nucleic acids, gen'l
- AMBER (Kollman) proteins and nucleic acids +
- BIO+ (Karplus) CHARMM; nucleic acids
- MacroModel (Still) biopolymers, general; good
- MMFF94 (Merck Pharm.) general; newer; good
- Sybyl in Alchemy2000, general (very poor).





## Existing Force Fields

- **AMBER (Assisted Model Building with Energy Refinement)**
  - Parameterized specifically for proteins and nucleic acids.
  - Uses only 5 bonding and non-bonding terms along with a sophisticated electrostatic treatment.
  - No cross terms are included.
  - Results can be very good for proteins and nucleic acids, less so for other systems.
- **CHARMM (Chemistry at Harvard Macromolecular Mechanics)**
  - Originally devised for proteins and nucleic acids.
  - Now used for a range of macromolecules, molecular dynamics, solvation, crystal packing, vibrational analysis and QM/MM studies.
  - Uses 5 valence terms, one of which is electrostatic term.



## Existing Force Fields

- **GROMOS (Gronigen molecular simulation)**
  - Popular for predicting the dynamical motion of molecules and bulk liquids.
  - Also used for modeling biomolecules.
  - Uses 5 valence terms, one of which is an electrostatic term.
- **MM1, 2, 3, 4**
  - General purpose force fields for (mono-functional) organic molecules.
  - MM2 was parameterized for a lot of functional groups.
  - MM3 is probably one of the most accurate ways of modeling hydrocarbons.
  - MM4 is very new and little is known about its performance.
  - The force fields use 5 to 6 valence terms, one of which is an electrostatic term and one to nine cross terms.



## Existing Force Fields

### ■ MMFF (Merck Molecular Force Field)

- General purpose force fields mainly for organic molecules.
- MMFF94 was originally designed for molecular dynamics simulations but is also widely used for geometry optimization.
- Uses 5 valence terms, one of which is an electrostatic term and one cross term.
- MMFF was parameterized based on high level *ab initio* calculations.

### ■ OPLS (Optimized Potential for Liquid Simulations)

- Designed for modeling bulk liquids.
- Has been extensively used for modeling the molecular dynamics of biomolecules.
- Uses 5 valence terms, one of which is an electrostatic term but no cross terms.



## Existing Force Fields

### ■ Tripos (SYBYL force field)

- Designed for modeling organic and biomolecules.
- Often used for CoMFA analysis (QSAR method).
- Uses 5 valence terms, one of which is an electrostatic term

### ■ MOMECC

- A force field for describing transition metals coordination compounds.
- Originally parameterized to use 4 valence terms but not an electrostatic term.
- Metal-ligand interactions consist of bond stretch only.
- Coordination sphere is maintained by non-bonding interactions between ligands.
- Work reasonable well for octahedrally coordinated compounds



## Existing Force Fields

### ■ UFF (Universal Force Field)

- Designed for coverage of the entire periodic table.
- All force field parameters are atomic based.
- Parameters for specific interactions are derived from atomic parameters based on a series of mixing rules, thereby addressing the problem of parameter explosion.
- Uses 4 valence terms but not an electrostatic term.

### ■ YATI

- Designed for the accurate representation of non-bonded interactions.
- Most often used for modeling interactions between biomolecules and small substrate molecules.



## Existing Force Fields

### ■ CVFF (Consistent Valence Force Field)

- Parameterized for small organic (amides, carboxylic acids, etc.) crystals and gas phase structures.
- Handles peptides, proteins, and a wide range of organic systems.
- Primarily intended for studies of structures and binding energies, although it predicts vibrational frequencies and conformational energies reasonably well.



## Existing Force Fields

- **CFF, CFF91, CFF95 (Consistent Force Field)**
  - Parameterized (based on quantum mechanics calculations and molecular simulations) for acetals, acids, alcohols, alkanes, alkenes, amides, amines, aromatics, esters, and ethers.
  - Useful for predicting:
    - Small molecules: gas-phase geometries, vibrational frequencies, conformational energies, torsion barriers, crystal structures.
    - Liquids: cohesive energy densities; for crystals: lattice parameters, RMS atomic coordinates, sublimation energies.
    - Macromolecules: protein crystal structures, polycarbonates and polysaccharides (CFF91,95).



## Existing Force Fields

- **Dreiding**
  - Parameters derived by a rule based approach.
  - Good coverage of the periodic table.
  - Good coverage for organic, biological and main-group inorganic molecules.
  - Only moderately accurate for geometries, conformational energies, intermolecular binding energies, and crystal packing.



## Molecular Mechanics Programs

- ↗ *HyperChem*
- ↗ *Alchemy2000*
- ↗ *Amber*
- ↗ *Charmm*
- ↗ *Spartan*
- ↗ and several others




## Uses of “Steric Energy”

- ↗ “Steric energy” has NO physical meaning, and it is defined differently in different programs
- ↗ Therefore it CAN NOT be used to compare structures calculated by different programs
- ↗ Its use is limited to comparing ISOMERIC structures having the SAME number and kinds of bonds (conformers, stereoisomers).



## Successes of Molecular Mechanics Calculations

- ↗ Calculations are very fast
- ↗ Geometry optimizations of small to medium- size molecules can be accomplished on a pc
- ↗ Conformations of macromolecules can be calculated using workstations or parallel processing computers.
- ↗ Reasonable geometries are usually obtained:
- ↗ Calculated energies are usually quite good:
- ↗ Provides input structure for more involved calculations (molecular orbital methods).



## Limitations of Molecular Mechanics

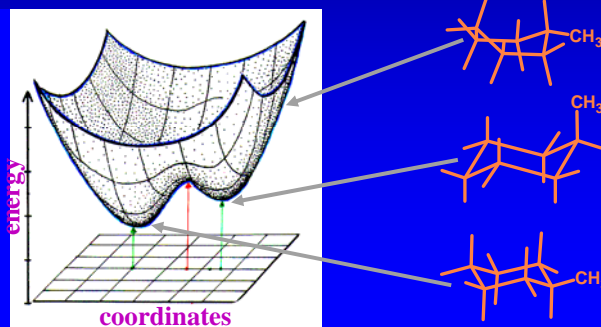
- ↗ The calculations do not account for electrons!  
Orbital interactions are ignored!
- ↗ The selection of “atom type” is crucial to the computational result:
  - ↗ e.g., AMBER has 5 types of Oxygen: carbonyl , alcohol, acid, ester/ether, water (see next slide)
  - ↗ No consideration is given to the importance of delocalized  $\pi$  electron systems
- ↗ Only ground states considered...not T.S. or \*

## Uses of Molecular Mechanics

- Obtaining a reasonably good geometry (in structures where pi electrons are not involved).
- As a starting point for further calculations, such as semi-empirical, *ab initio*, or density functional.
- Searching the energy surface for minimum energy conformations (it is usually too expensive to do this using MO methods).

## Force Field and Potential Energy Surface

- A force field defines for each molecule a unique PES.
- Each point on the PES represents a molecular conformation characterized by its structure and energy.
- Energy is a function of the coordinates.
- Coordinates are function of the energy.

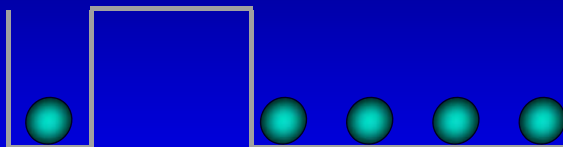


## Moving on (Sampling) the PES

- Each point on the PES is represents a molecular conformation characterized by its structure and energy.
- By sampling the PES we can derive molecular properties.
- Sampling energy minima only (energy minimization) will lead to molecular properties reflecting the enthalpy only.
- Sampling the entire PES (molecular simulations) will lead to molecular properties reflecting the free energy.
- In both cases, molecular properties will be derived from the PES.

## Deriving Molecular Properties from the PES

- Any molecular property is calculated by averaging the properties of all points (conformations) on the PES. The averaging weights are given by the relative energies of the conformations.

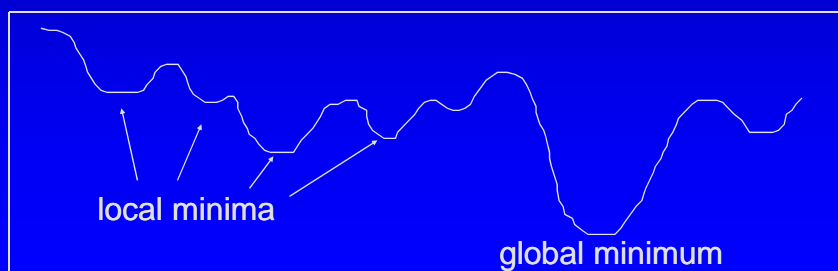


- **Energy minimization**
  - The two wells have equal enthalpies and thus contribute equally to each molecular property
- **Molecular Dynamics**
  - The two wells have different entropies and consequently free energies. The wider well contributes more to molecular properties

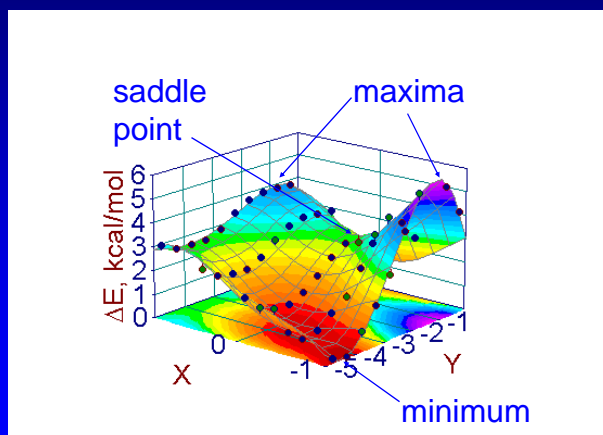


## Energy Minimization

- Local minimum vs global minimum
- Many local minima; only ONE global minimum
- Methods: Newton-Raphson (block diagonal), steepest descent, conjugate gradient, others.



## Potential Energy Surface





## PES & Energy Minimization

First, some caveats:

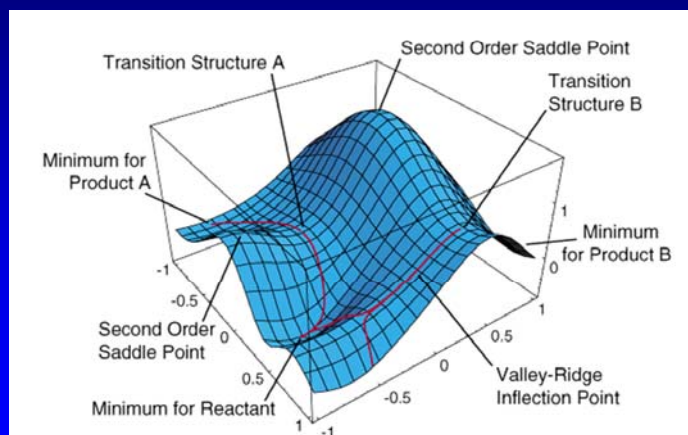
- extrema (stationary points) are located by most methods; this includes maxima, minima, and saddle points.
- among the minima, local minima are found, not necessarily the global minimum.
- with shallow minima (flat PES), a lot of cpu time can be spent seeking the lowest energy structure.



## Caveats about Minimum Energy Structures

- What does the global minimum energy structure mean?
- Does reaction/interaction of interest necessarily occur *via* the lowest energy conformation?
- What other low energy conformations are available? (Boltzmann distribution and probability/entropy considerations may be important).

## Potential Energy Surface



## 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 negative 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**, and higher order **Saddle Points**.

$$G = \left( \frac{\partial V}{\partial r} \right)_{r=r_e} = 0,$$



## PES Terms...

- In order to distinguish among the latter, one must examine the second derivatives of the PES with respect to geometry; the matrix of these is termed a **Hessian** (or force) matrix.
- Diagonalization of this matrix yields **Eigenvectors** which are normal modes of vibration; the **Eigenvalues** are proportional to the square of the vibrational frequency.

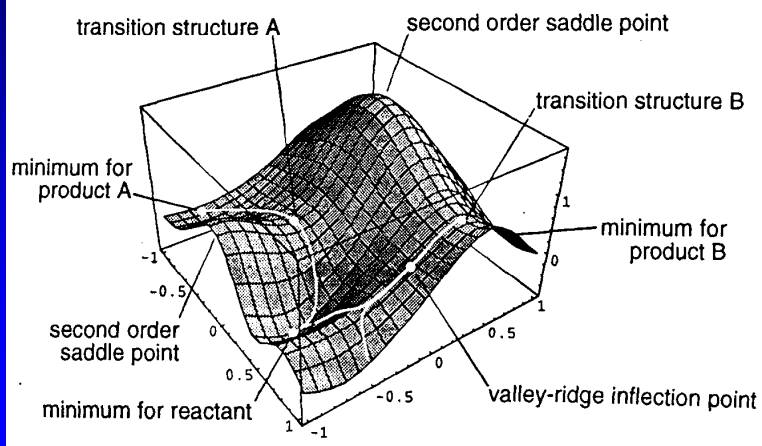
$$F = \left( \frac{\partial^2 V}{\partial r^2} \right)_{r=r_e}$$



## Sign of 2nd Derivatives

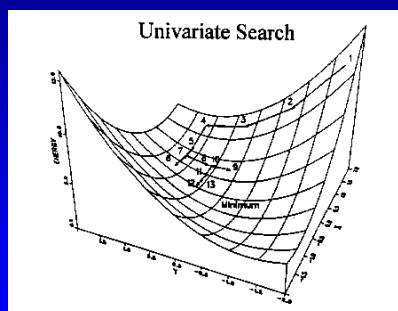
- The sign of the second derivative can be used to distinguish between **Maxima** and **Minima** on the PES
- **Minima** on the PES have only positive eigenvalues (vibrational frequencies)
- **Maxima** or **Saddle Points** (maximum in one direction but minimum in other directions) have one or more negative (*imaginary*) frequencies.

## Potential Energy Surface



## 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.



# Steepest Descent Method

- Simplest method in use
- Follows most negative gradient (max. force)
- Fastest method from a poor starting geometry
- Converges slowly near the energy minimum
- Can skip back and forth across a minimum.

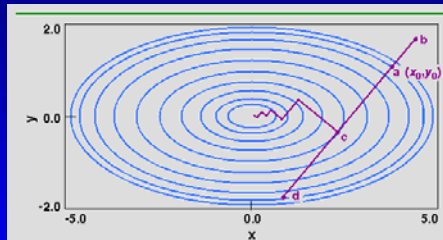


Figure 4-2. Energy Surface for Eq. 4-1

The derivative vector from the initial point  $a(x_0, y_0)$  defines the line search direction. Note that the derivative vector does not point directly toward the minimum. Compare this representation with that in Figure 4-3, where the line  $(b-a-c-d)$  is searched in one dimension for the minimum. Note that the minimum (point  $c$ ) occurs precisely at the point where the derivative vector is tangent to the energy contours, which implies that the subsequent derivative vectors are orthogonal to the previous derivatives.

# Steepest Descent Method

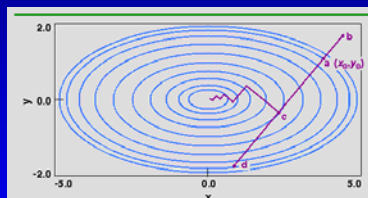


Figure 4-2. Energy Surface for Eq. 4-1

The derivative vector from the initial point  $a(x_0, y_0)$  defines the line search direction. Note that the derivative vector does not point directly toward the minimum. Compare this representation with that in Figure 4-3, where the line  $(b-a-c-d)$  is searched in one dimension for the minimum. Note that the minimum (point  $c$ ) occurs precisely at the point where the derivative vector is tangent to the energy contours, which implies that the subsequent derivative vectors are orthogonal to the previous derivatives.

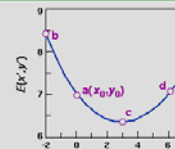
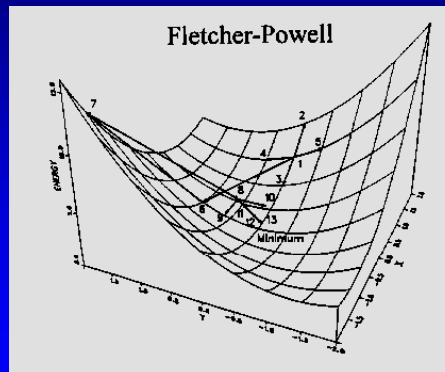


Figure 4-3. Cross Section of the Energy Surface as Defined by the Intersection of the Line Search Path in Figure 4-2 with the Energy Surface

The independent variable  $\alpha$  is a one-dimensional parameter that is adjusted so as to minimize the value of the function  $E(x', y')$ , where  $x'$  and  $y'$  are parameterized in terms of  $\alpha$  in Eq. 4-3. The point  $(a)$  corresponds to the initial point (when  $\alpha$  is 0), and point  $c$  is the local one-dimensional minimum. Points  $b$  and  $d$ , along with  $a$ , bound the minimum and form the basis for an iterative search for the minimum.

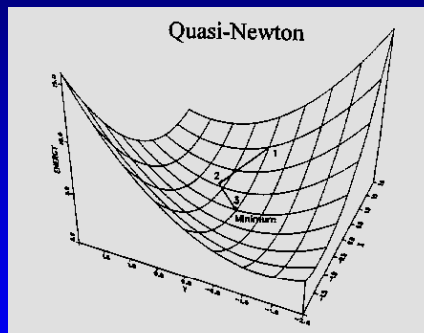
## Conjugate Gradient Method

- To simplicity of steepest descent method adds 'history' to implicitly gather 2nd derivative information to guide the search.
- Variations on this procedure are the Fletcher-Reeves, the Davidon- Fletcher-Powell and the Polak-Ribiere methods.



## Second Derivative Methods

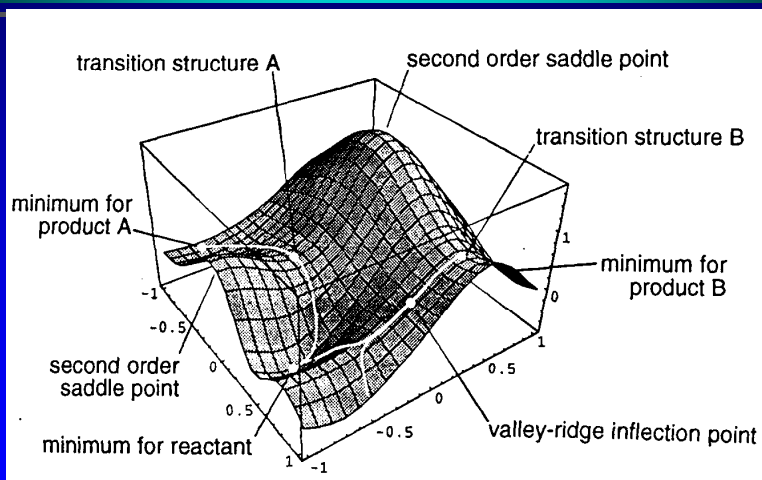
- The 2nd derivative of the energy with respect to  $x, y, z$  [the Hessian] determines the pathway.
- Computationally more involved, but generally fast and reliable, esp. near the minimum.
- Quasi-Newton, Newton-Raphson, block diagonal Newton-Raphson



## Transition State Modeling

- A Transition State is a stationary point for which the second derivative of the energy with respect to the reaction coordinate is negative, but second derivatives in other directions are positive.
- In other words, the TS is the highest point along the lowest energy pathway between reactants and products.

## Potential Energy Surface





## Transition States; Why Difficult?

- Reactants and products are well defined molecular entities; Transition States are not.
- It is thought that TS's exhibit elongated bonds, partial bonding, and may have some aspects of excited states associated with them.
- TS's cannot be observed experimentally; therefore no parameters can be devised for modeling them.

## TS Modeling Difficulties...

- Mathematically, there is less attention paid to saddle points than to minima, so there are fewer algorithms available to locate them.
- It is generally thought that the PES in the vicinity of the TS is 'flatter' than the surface near a minimum, therefore it may be more difficult to predict the structure of a transition state accurately. There may not exist a unique T.S. structure.

## Do all Reactions have a TS?

➤ No!!! There are numerous examples of barrierless reaction pathways:

➤ combination of radicals



➤ addition of radicals to alkenes



➤ gas phase addition of ions to neutral molecules

## Do all the paths lead to Rome?

