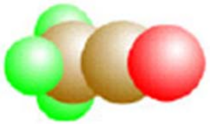
*Applied Statistical Mechanics*
*Lecture Note - 6*



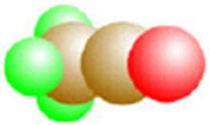
# Molecular Mechanics
# - Force Field Method

고려대학교
화공생명공학과
강정원

# Introduction

- Force Field Method vs. Electronic Structure Method
  - Force field method : based on Molecular Mechanics
  - Electronic structure method : based on Quantum Mechanics

# Force Field Method



- Problem
  - Calculating energy for given structure
  - Finding stable geometry of molecules
    - Energy optimum of saddle point
- Molecules are modeled as atoms held together with bonds
  - “Ball and spring model “
- Bypassing the electronic Schrödinger equation
- Quantum effects of nuclear motions are neglected
- The atom is treated by classical mechanics → Newton’s second law of motion

# Force Field Methods

- Validation of FF methods
  - Molecules tend to be composed of units which are structurally similar in different molecules
    - Ex) C-H bond
      - bond length : 1.06 – 1.10 A
      - stretch vibrations : 2900 – 3300 /cm
    - Heat of formation for $CH_3 – (CH2)_n – CH_3$ molecules
      - Almost straight line when plotted against n
- Molecules are composed of structural units

  → "Functional groups"

# Example MM2 atom types

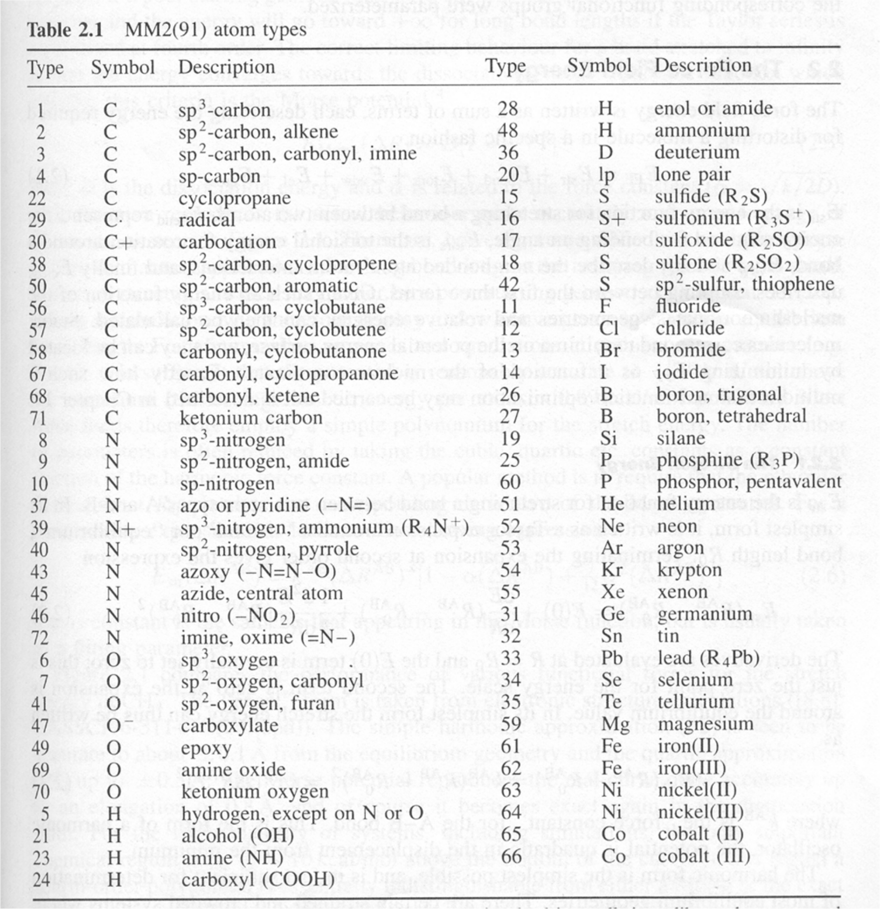


**Table 2.1** MM2(91) atom types

| Type | Symbol | Description | Type | Symbol | Description |
|---|---|---|---|---|---|
| 1 | C | $sp^3$-carbon | 28 | H | enol or amide |
| 2 | C | $sp^2$-carbon, alkene | 48 | H | ammonium |
| 3 | C | $sp^2$-carbon, carbonyl, imine | 36 | D | deuterium |
| 4 | C | sp-carbon | 20 | lp | lone pair |
| 22 | C | cyclopropane | 15 | S | sulfide ($R_2S$) |
| 29 | C· | radical | 16 | S+ | sulfonium ($R_3S^+$) |
| 30 | C+ | carbocation | 17 | S | sulfoxide ($R_2SO$) |
| 38 | C | $sp^2$-carbon, cyclopropene | 18 | S | sulfone ($R_2SO_2$) |
| 50 | C | $sp^2$-carbon, aromatic | 42 | S | $sp^2$-sulfur, thiophene |
| 56 | C | $sp^3$-carbon, cyclobutane | 11 | F | fluoride |
| 57 | C | $sp^2$-carbon, cyclobutene | 12 | Cl | chloride |
| 58 | C | carbonyl, cyclobutanone | 13 | Br | bromide |
| 67 | C | carbonyl, cyclopropanone | 14 | I | iodide |
| 68 | C | carbonyl, ketene | 26 | B | boron, trigonal |
| 71 | C | ketonium carbon | 27 | B | boron, tetrahedral |
| 8 | N | $sp^3$-nitrogen | 19 | Si | silane |
| 9 | N | $sp^2$-nitrogen, amide | 25 | P | phosphine ($R_3P$) |
| 10 | N | sp-nitrogen | 60 | P | phosphor, pentavalent |
| 37 | N | azo or pyridine (–N=) | 51 | He | helium |
| 39 | N+ | $sp^3$-nitrogen, ammonium ($R_4N^+$) | 52 | Ne | neon |
| 40 | N | $sp^2$-nitrogen, pyrrole | 53 | Ar | argon |
| 43 | N | azoxy (–N=N–O) | 54 | Kr | krypton |
| 45 | N | azide, central atom | 55 | Xe | xenon |
| 46 | N | nitro ($-NO_2$) | 31 | Ge | germanium |
| 72 | N | imine, oxime (=N–) | 32 | Sn | tin |
| 6 | O | $sp^3$-oxygen | 33 | Pb | lead ($R_4Pb$) |
| 7 | O | $sp^2$-oxygen, carbonyl | 34 | Se | selenium |
| 41 | O | $sp^2$-oxygen, furan | 35 | Te | tellurium |
| 47 | $O^-$ | carboxylate | 59 | Mg | magnesium |
| 49 | O | epoxy | 61 | Fe | iron(II) |
| 69 | O | amine oxide | 62 | Fe | iron(III) |
| 70 | O | ketonium oxygen | 63 | Ni | nickel(II) |
| 5 | H | hydrogen, except on N or O | 64 | Ni | nickel(III) |
| 21 | H | alcohol (OH) | 65 | Co | cobalt (II) |
| 23 | H | amine (NH) | 66 | Co | cobalt (III) |
| 24 | H | carboxyl (COOH) | | | |

# The Force Field Energy

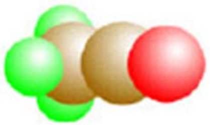
- Expressed as a sum of terms
  - ☐ $E_{str}$ : The stretch energy
  - ☐ $E_{bend}$ : The bending energy
  - ☐ $E_{tor}$ : The torsion energy
  - ☐ $E_{vdw}$ : The van der Waals energy
  - ☐ $E_{el}$ : The electro static energy
  - ☐ $E_{cross}$ : coupling between the first three terms

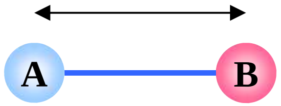
$E_{str}$, $E_{bend}$, $E_{tor}$ → **Bonded interactions**

$E_{vdw}$, $E_{el}$ → **Nonbonded atom-atom interaction**

$$E_{FF} = E_{str} + E_{bend} + E_{tor} + E_{vdw} + E_{el} + E_{cross}$$

# The stretch energy





- $E_{str}$ : The energy function for stretching a bond between tow atom type A and B
- Equilibrium bond length → Minimum energy
- Taylor series expansion around equilibrium bond length

**Set to 0** **0 at minimum energy**

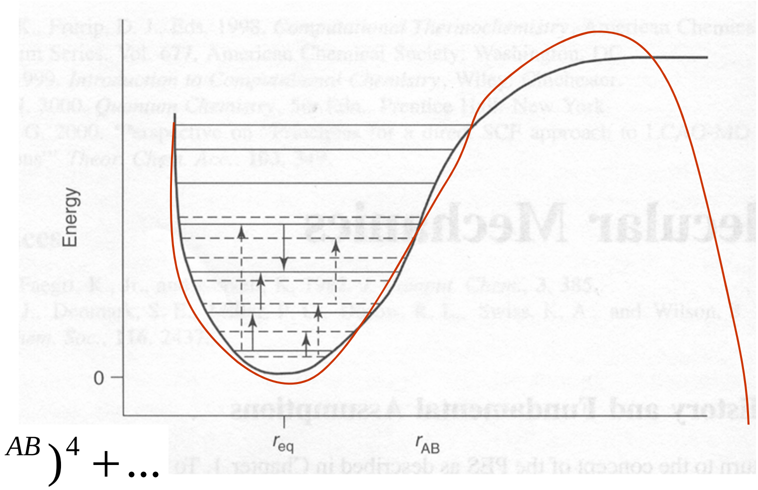
$$E_{str}(R^{AB}-R_0^{AB}) = E(0) + \frac{dE}{dR}(R^{AB}-R_0^{AB}) + \frac{1}{2}\frac{d^2E}{dR^2}(R^{AB}-R_0^{AB})^2 + \ldots$$

**Simplest form**
**: Harmonic Oscillator**

$$E_{str}(R^{AB}-R_0^{AB}) = k^{AB}(R^{AB}-R_0^{AB})^2 = k^{AB}(\Delta R^{AB})^2$$

# The stretch energy

- The harmonic form is the simplest possible form
- When the bond is stretched to longer r , the predicted energy is not realistic
- Polynomial expansion



$$E_{str}(\Delta R^{AB}) = k_2^{AB}(\Delta R^{AB})^2 + k_3^{AB}(\Delta R^{AB})^3 + k_4^{AB}(\Delta R^{AB})^4 + \ldots$$

  - More parameters
  - The limiting behavior is not correct for some cases ( 3rd order, 5th order,…)
  - Special care needed for optimization (negative energy for long distance )

# The stretch energy

- The Morse Potential

$$E_{str}(\Delta R^{AB}) = D[1-e^{\alpha\Delta R}]^2$$

$$\alpha = \sqrt{k/2D}$$

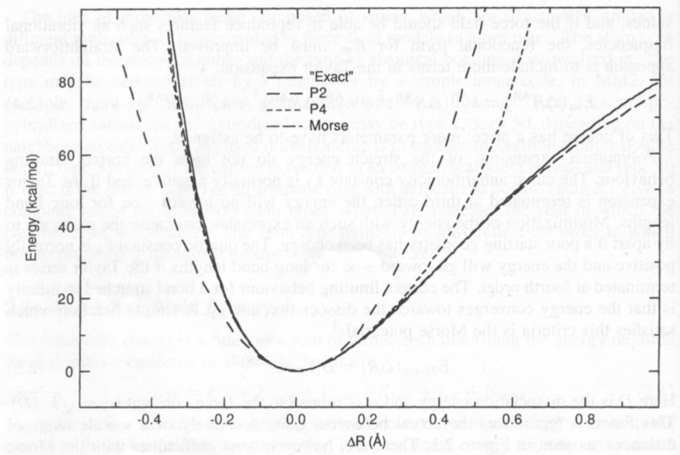


**Figure 2.1** The stretch energy for $CH_4$

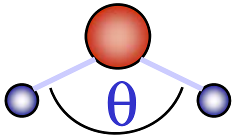
  - ☐ *D*: Dissociation energy
  - ☐ Accurate actual behavior
  - ☐ Problem
    - More computation time evaluating exponential term
    - Starting from poor geometry, slow convergence

- Popular method : n$^{th}$ order expansion of the Morse Potential

$$E_{str}(\Delta R^{AB}) = k^{AB}(\Delta R^{AB})^2[1-\alpha(\Delta R^{AB}) + \frac{7}{12}\alpha^2(\Delta R^{AB})^2]$$

# The Bending Energy

- $E_{bend}$ : The energy required for bending an angle formed by three atoms A-B-C



- Harmonic Approximation

$$E_{bend}(\theta^{ABC}-\theta_0^{ABC})=k^{ABC}(\theta^{ABC}-\theta_0^{ABC})^2$$

- Improvement can be observed when more terms are included
  - ☐ Adjusting higher order term to fixed fraction
- For most applications, simple harmonic approximation is quite adequate
  - ☐ MM3 force field : 6th term

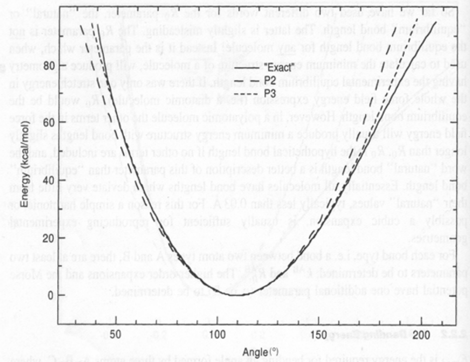


Figure 2.3 The bending energy for $CH_4$

# The bending energy

- Angels where the central atom is di- or tri-valent (ethers, alcohols, sulfiteds, amines), represents a special problem
  - an angle of 180 degree → energy maximum
  - at least order of three
- Refinement over a simple harmonic potential clearly improve the overall performance.
- They have little advantage in the chemically important region ( 10 kcal/mol above minimum)

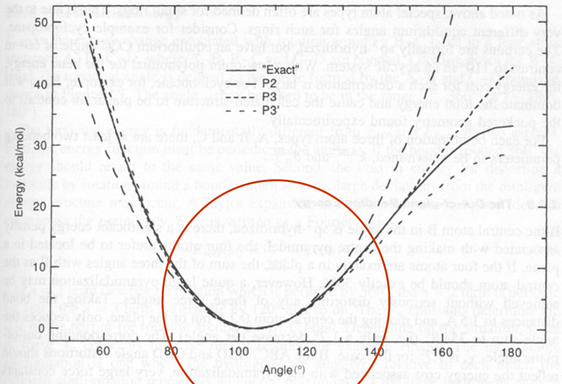


Figure 2.5 The bending energy for $H_2O$

# The out-of-plane bending energy

- *sp*$^2$-hybridized atoms (ABCD)
  - ☐ there is a significant energy penalty associated with making the center pyramidal
  - ☐ ABD, ABD, CBD angle distortion should reflect the energy cost associated with pyramidization

$$E_{oop}(\chi_B) = k^B(\chi_B)^2$$

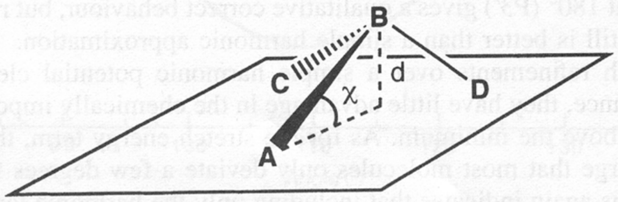
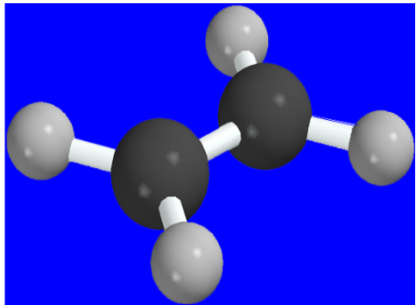


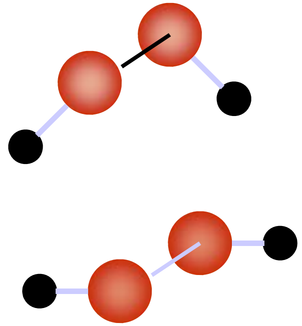
**Figure 2.6** Out-of-plane variable definitions

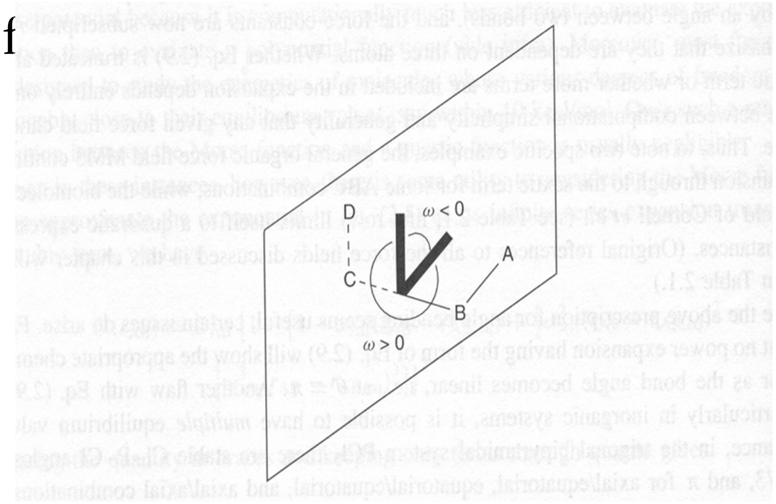


# The torsion energy

- Angle of rotation around B-C bond for four atoms sequence A-B-C-D
- Difference between stretch and bending energy
  - The energy function must be periodic win the angle $\omega$
  - The cost of energy for distortion is of low
    - Large deviation from minimum can occur
- Fourier series expansion

$$E_{tors}(\omega) = \sum_{n=1} V_n \cos(n\omega)$$

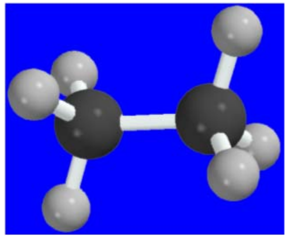




# The torsion energy

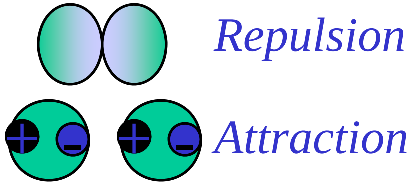
$$E_{tors}(\omega) = \sum_{n=1} V_n \cos(n\omega)$$

- Depending on the situation some of $V_n$ terms are set to 0
  - $n=1$ : periodic by 360 degree
  - $n=2$ : periodic by 170 degree
  - $n=3$ : periodic by 120 degree
  - Ethane : three minima and three maxima
    - $n = 3,6,9, \ldots$ can have $V_n$



# The van der Waals energy

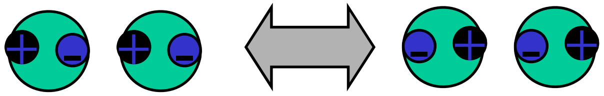
- $E_{vdw}$ : energy describing the repulsion and attraction between atoms : non-bonded energy
  - Interaction energy not related to electrostatic energy due to atomic charges
  - Repulsion and attraction
    - Small distance, very repulsive → overlap of electron cloud
    - Intermediate distance, slight attraction → electron correlation
      - motion of electrons create temporarily induced dipole moment



# Van der Waals Attraction

- 1930, London
  - ☐ "Dispersion" or "London" force
  - ☐ Correlation of electronic fluctations
  - ☐ Explained attraction as induced dipole interaction

$$E_{vdw}^{attr}(R) \propto \frac{C}{R^6} + O(R^{-8})$$

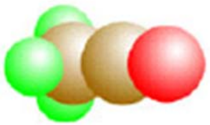


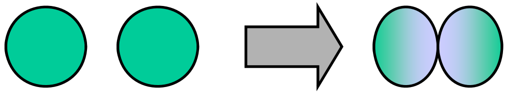
$$\text{induced dipole - dipole interaction} \propto \frac{1}{R^6}$$

$$\text{induced dipole - quadrupole interaction} \propto \frac{1}{R^8}$$

$$\text{induced quadrupole - quadrupole interaction} \propto \frac{1}{R^{10}}$$

# Van der Waals Repulsion





- Overlap of electron cloud
- Theory provide little guidance on the form of the model
- Two popular treatment
  - ☐ Inverse power
    - Typically n = 9 -12

$$E_{vdw}^{rep}(R) \propto \frac{C}{R^n}$$

  - ☐ Exponential
    - Two parameters A, B

$$E_{vdw}^{rep}(R) \propto Ae^{-BR}$$

# Van der Waals Energy

- Repulsion + Atrraction gives two model
  - Lennard-Jones potential

$$E_{LJ}(R) = \varepsilon\left[\left(\frac{R_0}{R}\right)^{12} - 2\left(\frac{R_0}{R}\right)^{6}\right]$$

  - Exp-6 potential
    - Also known as “Buckingham” or “Hill” type potential

$$E_{vdw}(R) = Ae^{-BR} - \frac{C}{R^6} = \varepsilon\left[\frac{6}{\alpha-6}e^{\alpha(1-R/R_0)} - \frac{6}{\alpha-6}\left(\frac{R_0}{R}\right)^{6}\right]$$

# Comparison

**Morse Potential**

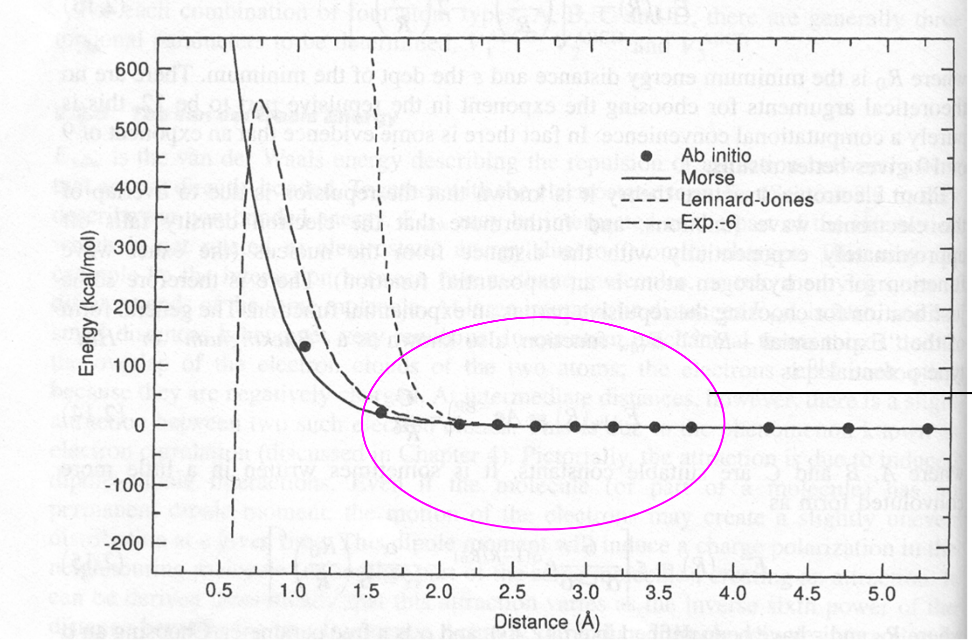
$$E_{str}(\Delta R^{AB}) = D[1-e^{\alpha\Delta R}]^2$$

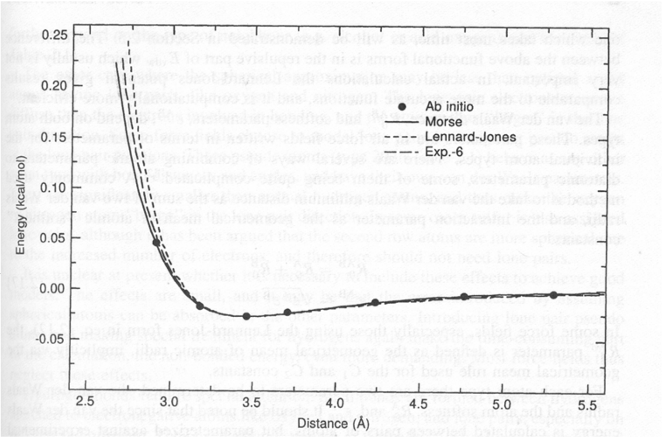
$$\alpha = \sqrt{k/2D}$$

**Problems**

**Inversion**

**Overestimating repulsion**





# Why LJ potential is preferred ?

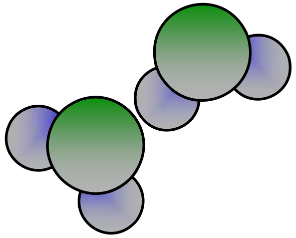
- Multiplications are much faster than exponential calculation
- Parameters are meaningful than the other models
- Diatomic parameters

$$\Delta R_0^{AB} = \frac{1}{2}(R_0^A + R_0^B)$$

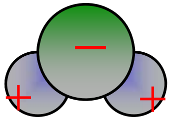
$$\varepsilon^{AB} = \sqrt{\varepsilon^A + \varepsilon^B}$$

# The electrostatic Energy

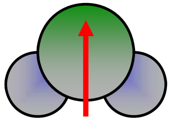
- Iternal distribution of electrons
  - positive and negative part of molecule
  - long range force than van der Waals

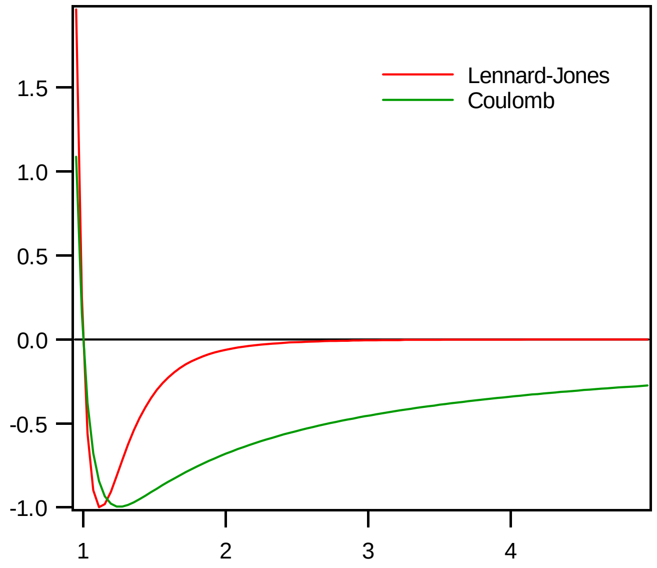


- Two modeling approaches
  - Point charges



  - Bond Dipole Description





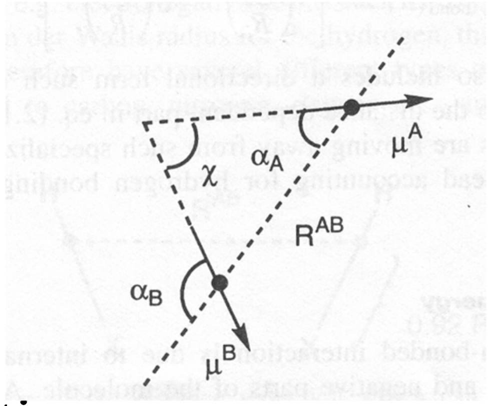
# Point charge method

- Assign Columbic charges to several points of molecules
- Total charge sum to charges on the molecule
- Atomic charges are treated as fitting parameters
  - Obtained from electrostatic potential calculated by electronic structure method (QM)

$$E_{el}(R^{AB}) = \frac{Q_A Q_B}{\varepsilon R^{AB}}$$

# Bond Dipole description

- Interaction between two dipole



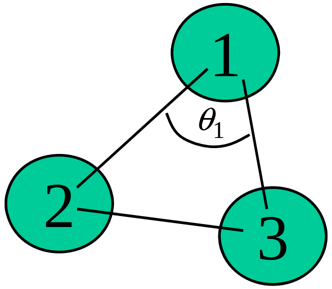
$$E_{el}(R^{AB}) = \frac{\mu_A \mu_B}{\varepsilon (R^{AB})^3} (\cos\chi - 3\cos\alpha_A \cos\alpha_B)$$

- MM2 and MM3 uses bond dipole description
- Point charge vs. Bond Dipole model
  - There is little difference if properly parameterized
  - The atomic charge model is easier to parameterize by fitting an electronic wave function → preferred by almost all force field

# Multibody interaction

- Unlike van der Waals interaction, the three body interaction is quite significant for polar species
- Two method
  - Explicit multibody interaction
    - Axilrod – Teller

$$u(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)=\frac{3}{2}\frac{E\alpha_1\alpha_2\alpha_3}{r_{12}^3 r_{23}^3 r_{13}^3}(3\cos\theta_1\cos\theta_2\cos\theta_3+1)$$



  - Atom Polarization
    - Electrostatic interaction = (Intrisic contribution) + (dipolar term arising from the other atomic charges)
    - Solved iterative self-consistent calculation

# Cross terms

- Bonds, angles and torsions are not isolated
  - They couple with one another
- Example
  - Stretch/bend coupling
  - Stretch/stretch coupling
  - Bend/bend coupling
  - Stretch/torsion coupling
  - Bend/torsion coupling
  - Bend/torsion/bend coupling …

# Small rings

- Small rings present a problem
  - their equilibrium angles are very different from those of their acyclic cousins
- Methods
  - Assign new atom types
  - Adding sufficient parameters in cross terms

# Conjugated systems

- Butadiene (C=C-C=C )
  - Same set of parameters are used for all carbon atoms
  - Bond length of terminal and central bonds are different ( 1.35 A and 1.47 A)
  - Delocalization of pi-electrons in the conjugated system
- Approaches
  - Identifying bond combination and use specialized parameters
  - Perform simple electronic structure calculation
    - Implemented in MM2 / MM3 (MMP2 and MMP3)
    - Electronic structure calculation method (Pariser-Pople-Parr (PPP) type) : Extended Hückel calculation
    - Requires additional second level of iteration in geometry optimization

***If the system of interest contains conjugation, a FF which uses the parameter replacement is chosen, the user should check that the proper bond length and reasonable rotation barrier !***

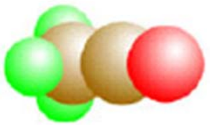
# Comparing Energies of Different Molecules

- The numerical value of force-field energy has no meaning !
  - Zero point energy has been chosen for convenienece
  - It is inconsequential for comparing energies of different conformation
  - $E_{FF}$ : “ steric energy”
- Heat of formation
  - Bond dissociation energy for each bond type
  - To achieve better fit, parameters may also be assigned to larger units (groups : $CH_3$- ,…)

$$\Delta H_f = E_{FF} + \sum_{bonds} \Delta H_{AB} + \sum_{groups} \Delta H_G$$

- MM2/MM3 attemped to parameterize heat of formation
- Other force fields are only concerned with producing geometries

# Force Field Parameterization



- Numerical Values of parameters
- Example : MM2 (71 atom types)
  - For one parameters at least 3-4 independent data are required
  - Require order of $10^7$ independent experimental data → Impossible
  - Rely on electronic structure calculation **("Class II" force field)**

| Term | Estimated | Actual | Calc. Basis |
|---|---|---|---|
| $E_{vdw}$ | 142 | 142 | 2*71 =142 |
| $E_{str}$ | 900 | 290 | (30*30/2)*2 |
| $E_{bend}$ | 27,000 | 824 | (30*30*30/2)*2 |
| $E_{tors}$ | 1,215,00 | 2,466 | (30*30*30*30/2)*3 |

# Force Field Parameterization

- There are large number of possible compounds for which there are no parameters
  - The situation is not as bad as it would appear
    - Although about 0.2 % of possible torsional constants have been parameterized,
    - About 20 % of 15 million known compound can be modeled by MM2

# Universal Force Field

- Many atom type, lack of sufficient reference data → Development of *Universal Force Field* (UFF)

- Derive di-, tri-, tetra- atomic parameters from atomic constant (Reduced parameter form)
  - Atomic properties : atom radii, ionization potential, electronegativity, polarizability, …
- In principle, capable of covering molecules composed of elecments from the whole periodic table
- Less accurate result presently …→ likely to be improved …

# Force Fields…

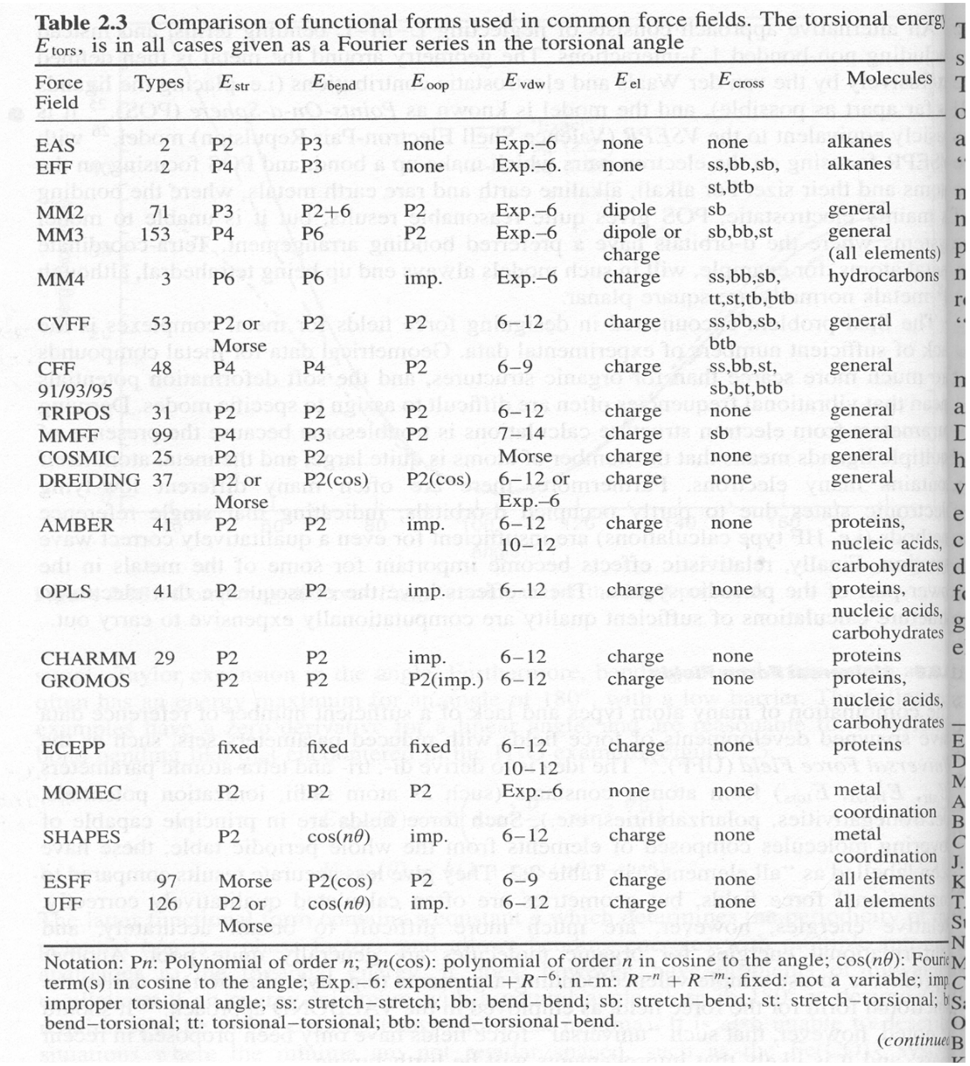


**Table 2.3** Comparison of functional forms used in common force fields. The torsional energ $E_{\text{tors}}$, is in all cases given as a Fourier series in the torsional angle

| Force Field | Types | $E_{\text{str}}$ | $E_{\text{bend}}$ | $E_{\text{oop}}$ | $E_{\text{vdw}}$ | $E_{\text{el}}$ | $E_{\text{cross}}$ | Molecules |
|---|---|---|---|---|---|---|---|---|
| EAS | 2 | P2 | P3 | none | Exp.–6 | none | none | alkanes |
| EFF | 2 | P4 | P3 | none | Exp.–6. | none | ss,bb,sb, st,btb | alkanes |
| MM2 | 71 | P3 | P2+6 | P2 | Exp.–6 | dipole | sb | general |
| MM3 | 153 | P4 | P6 | P2 | Exp.–6 | dipole or charge | sb,bb,st | general (all elements) |
| MM4 | 3 | P6 | P6 | imp. | Exp.–6 | charge | ss,bb,sb, tt,st,tb,btb | hydrocarbons |
| CVFF | 53 | P2 or Morse | P2 | P2 | 6–12 | charge | ss,bb,sb, btb | general |
| CFF 91/93/95 | 48 | P4 | P4 | P2 | 6–9 | charge | ss,bb,st, sb,bt,btb | general |
| TRIPOS | 31 | P2 | P2 | P2 | 6–12 | charge | none | general |
| MMFF | 99 | P4 | P3 | P2 | 7–14 | charge | sb | general |
| COSMIC | 25 | P2 | P2 | | Morse | charge | none | general |
| DREIDING | 37 | P2 or Morse | P2(cos) | P2(cos) | 6–12 or Exp.–6 | charge | none | general |
| AMBER | 41 | P2 | P2 | imp. | 6–12<br>10–12 | charge | none | proteins, nucleic acids, carbohydrates |
| OPLS | 41 | P2 | P2 | imp. | 6–12 | charge | none | proteins, nucleic acids, carbohydrates |
| CHARMM | 29 | P2 | P2 | imp. | 6–12 | charge | none | proteins |
| GROMOS | | P2 | P2 | P2(imp.) | 6–12 | charge | none | proteins, nucleic acids, carbohydrates |
| ECEPP | | fixed | fixed | fixed | 6–12<br>10–12 | charge | none | proteins |
| MOMEC | | P2 | P2 | P2 | Exp.–6 | none | none | metal coordination |
| SHAPES | | P2 | $\cos(n\theta)$ | imp. | 6–12 | charge | none | metal coordination |
| ESFF | 97 | Morse | P2(cos) | P2 | 6–9 | charge | none | all elements |
| UFF | 126 | P2 or Morse | $\cos(n\theta)$ | imp. | 6–12 | charge | none | all elements |

Notation: P*n*: Polynomial of order *n*; P*n*(cos): polynomial of order *n* in cosine to the angle; $\cos(n\theta)$: Fouri term(s) in cosine to the angle; Exp.–6: exponential $+R^{-6}$; n–m: $R^{-n}+R^{-m}$; fixed: not a variable; imp improper torsional angle; ss: stretch–stretch; bb: bend–bend; sb: stretch–bend; st: stretch–torsional; b bend–torsional; tt: torsional–torsional; btb: bend–torsional–bend.

(continue