Molecular Mechanics Force Field

The "mechanical" molecular model was developed out of a need to describe molecular structures and properties in as practical a manner as possible. The range of applicability of molecular mechanics includes:

- Molecules containing thousands of atoms.
- Organics, oligonucleotides, peptides, and saccharides (metallo-organics and inorganics in some cases).
- Vacuum, implicit, or explicit solvent environments.
- Ground state only.
- Thermodynamic and kinetic (via molecular dynamics) properties.

The great computational speed of molecular mechanics allows for its use in procedures such as molecular dynamics, conformational energy searching, and docking. All the procedures require large numbers of energy evaluations.

Molecular mechanics methods are based on the following principles:

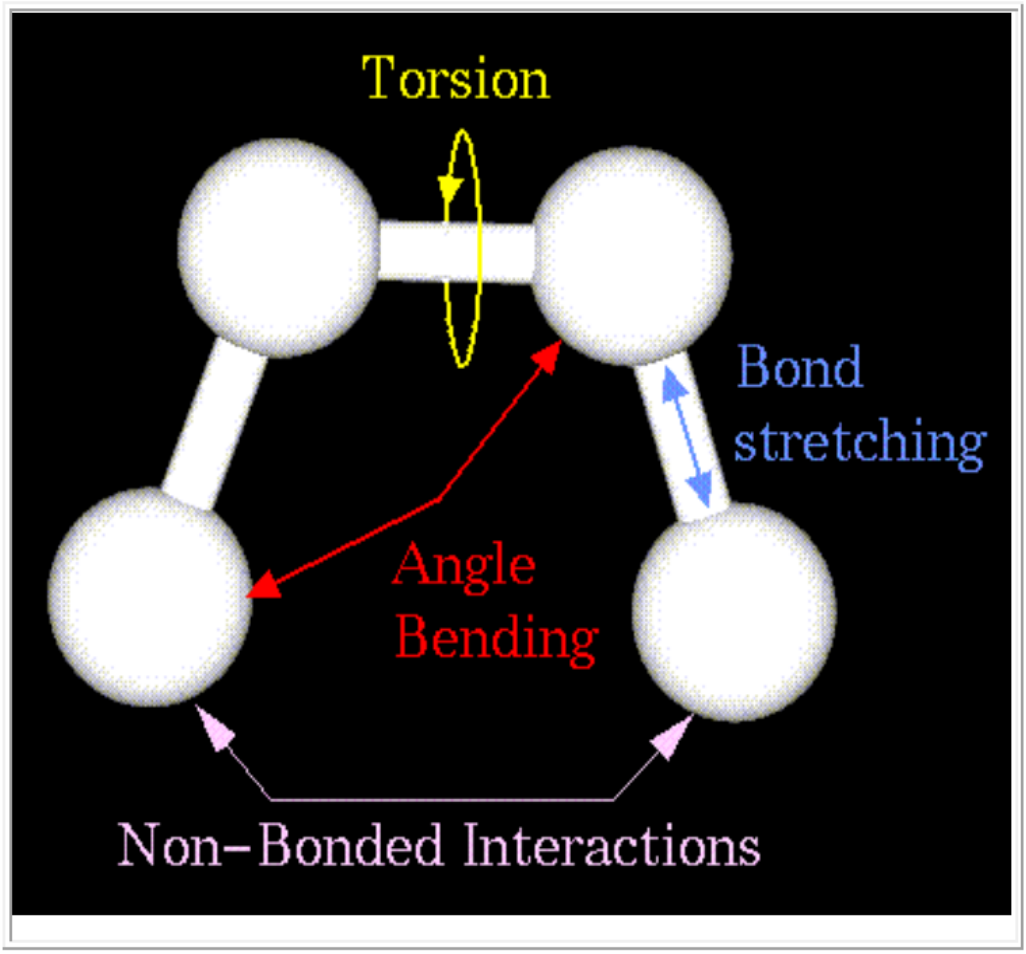
- Nuclei and electrons are lumped into atom-like particles.
- Atom-like particles are spherical (radii obtained from measurements or theory) and have a net charge (obtained from theory).
- Interactions are based on springs and classical potentials.
- Interactions must be preassigned to specific sets of atoms.

Interactions determine the **spatial distribution** of atom-like particles and their energies.

To define a force field one must specify not only the functional form but also the parameters (i.e.the various constants). Two force fields may use an identical functional form yet have very different parameters. A force field should be considered as a single entity; it is

not strictly correct to divide the energy into its individual components, let alone to take some of the parameters from one forcefield and mix them with parameters from another force field. The forcefields used in molecular modelling are primarily designed to reproduce structural properties but they can also be used to predict other properties, such as molecular spectra. However, molecular mechanics force fields can rarely predict spectra with great accuracy (although the more recent molecular. mechanics force fields are much better in this regard). A force field is generally designed to predict certain properties and will be parametrised accordingly. While it is useful to try to predict other quantities which have not been included in the parametrisation process it is not necessarily a failing if a force field is unable to do so. Transferability of the functional form and parameters is an important feature of a forcefield. Transferability means that the same set of parameters can be used to model a series of related molecules, rather than having to define a new set of parameters for each individual molecule. A concept that is common to most force fields is that of an atom type. When preparing the input for a quantum mechanics calculation it is usually necessary to specify the atomic numbers of the nuclei present, together with the geometry of the system and the overall charge and spin multiplicity. For a force field the overall charge and spin multiplicity are not explicitly required, but it is usually necessary to assign an atom type to each atom in the system. The atom type is more than just the atomic number of an atom; it usually con• tains information about its hybridisation state and sometimes the local environment. For example, it is necessary in most force fields to distinguish between sp3 - hybridised carbon atoms (which adopt a tetrahedral geometry), sp2-hybridised carbons (which are trigonal) and sp-hybridised carbons (which are linear).

The mechanical molecular model considers atoms as spheres and bonds as springs. The mathematics of spring deformation can be used to describe the ability of bonds to stretch, bend, and twist:



Non-bonded atoms (greater than two bonds apart) interact through van der Waals attraction, steric repulsion, and electrostatic attraction/repulsion. These properties are easiest to describe mathematically when atoms are considered as spheres of characteristic radii.

The object of molecular mechanics is to predict the energy associated with a given conformation of a molecule. However, molecular mechanics energies have no meaning as absolute quantities. Only differences in energy between two or more conformations have meaning. A simple molecular mechanics energy equation is given by:

Energy = Stretching Energy + Bending Energy + Torsion Energy + Non-Bonded Interaction Energy

- A **force field** refers to the form and parameters of mathematical functions used to describe the potential energy of a system of particles (typically molecules and atoms).
- calculates the molecular system's potential energy (E) in a given conformation as a sum of individual energy terms.
- where the components of the covalent and noncovalent contributions are given by the following summations:

$$E_{\text{noncovalent}} = E_{\text{electrostatic}} + E_{\text{van der Waals}}$$

- where the components of the covalent and noncovalent contributions are given by the following summations

$$E_{\text{covalent}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}}$$

$$E_{\text{noncovalent}} = E_{\text{electrostatic}} + E_{\text{van der Waals}}$$

- FF is a mathematical function which returns the energy of the system as a function of the conformation of the system.

These equations together with the data (parameters) required to describe the behavior of different kinds of atoms and bonds, is called a force-field. Many different kinds of force-fields have been developed over the years. Some include additional energy terms that describe other kinds of deformations. Some force-fields account for coupling between bending and stretching in adjacent bonds in order to improve the accuracy of the mechanical model.

$$\mathscr{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$
$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

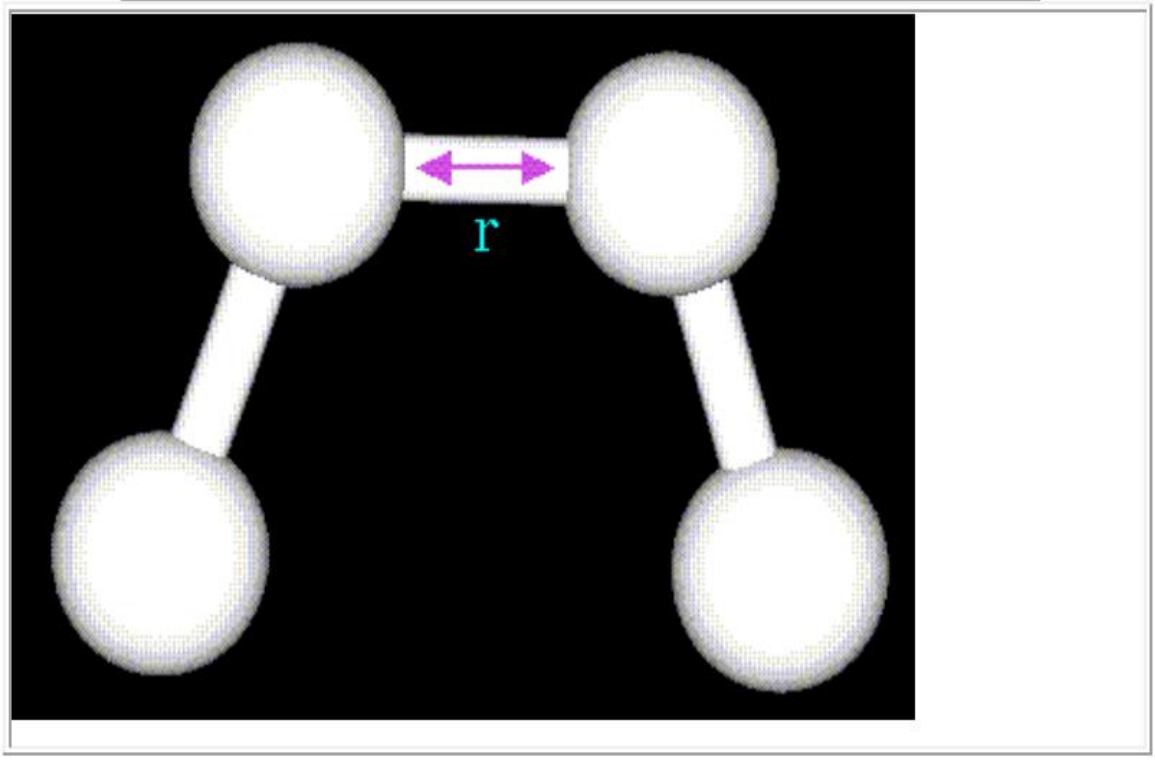
$\mathcal{V}(\mathbf{r}^N)$ Potential energy as a function of position r of N particles

- Reproduce the structural properties such as molecular spectra
- Transferability

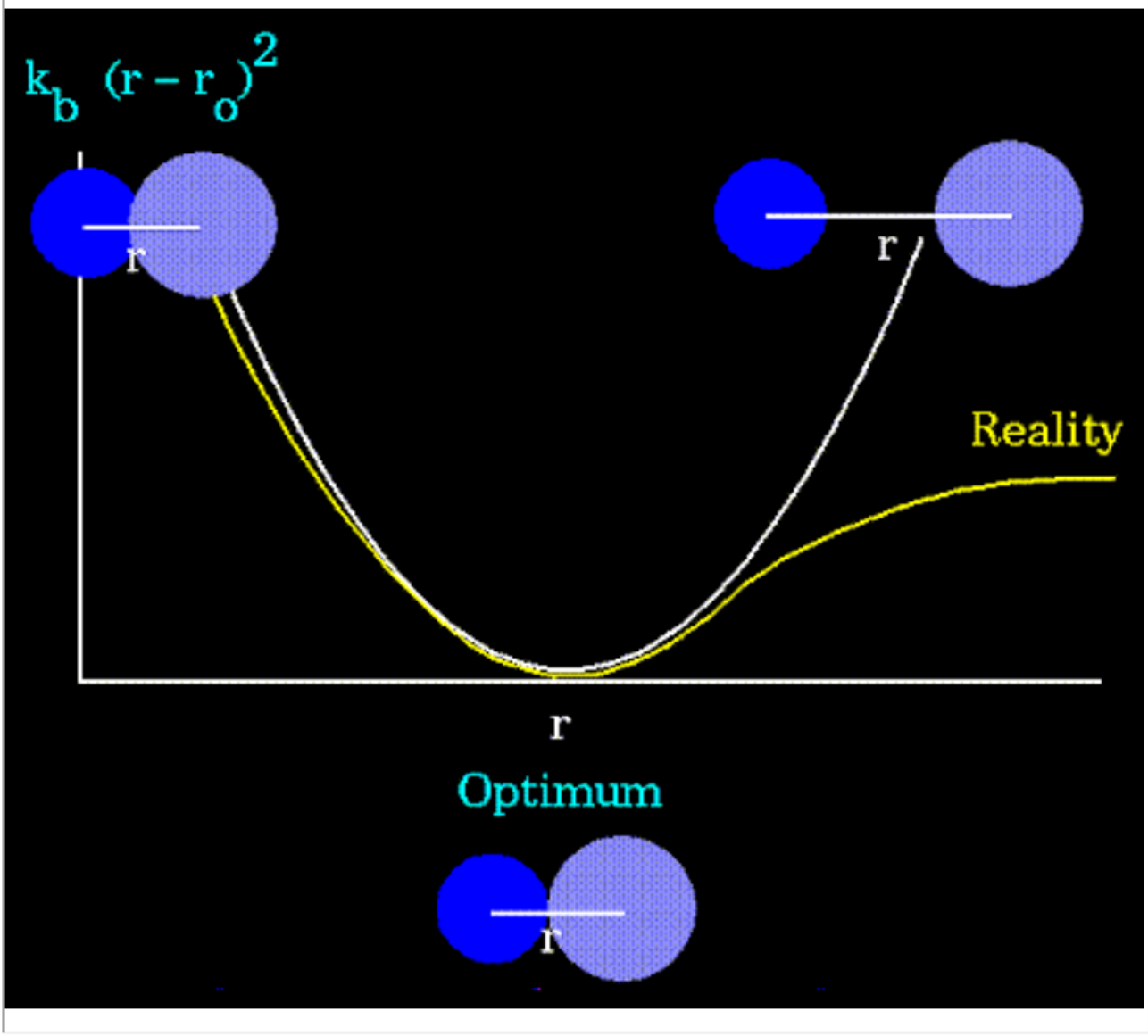
The mathematical form of the energy terms varies from force-field to force-field. The more common forms will be described.

Stretching Energy

$$E = \sum_{bonds} k_b (r - r_o)^2$$



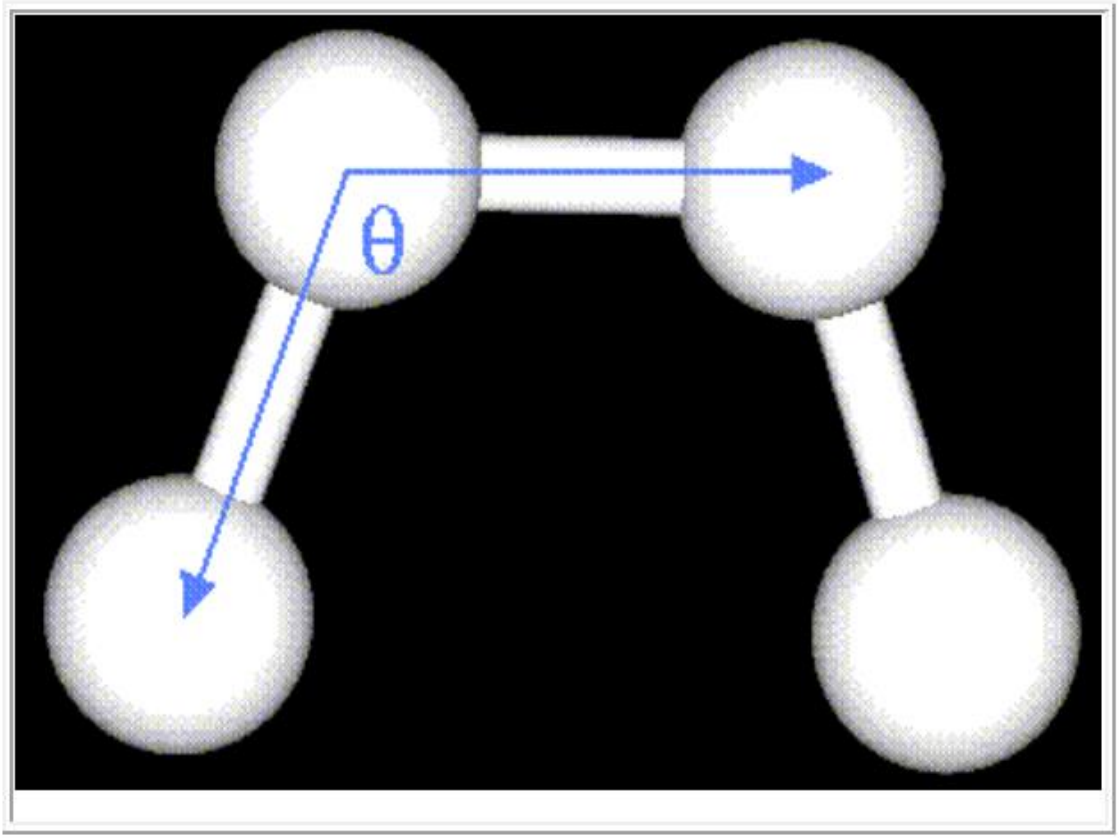
The stretching energy equation is based on Hooke's law. The "k*b*" parameter controls the stiffness of the bond spring, while "r*o*" defines its equilibrium length. Unique "k*b*" and "r*o*" parameters are assigned to each pair of bonded atoms based on their types (e.g. C-C, C-H, O-C, etc.). This equation estimates the energy associated with vibration about the equilibrium bond length. This is the equation of a parabola, as can be seen in the following plot:



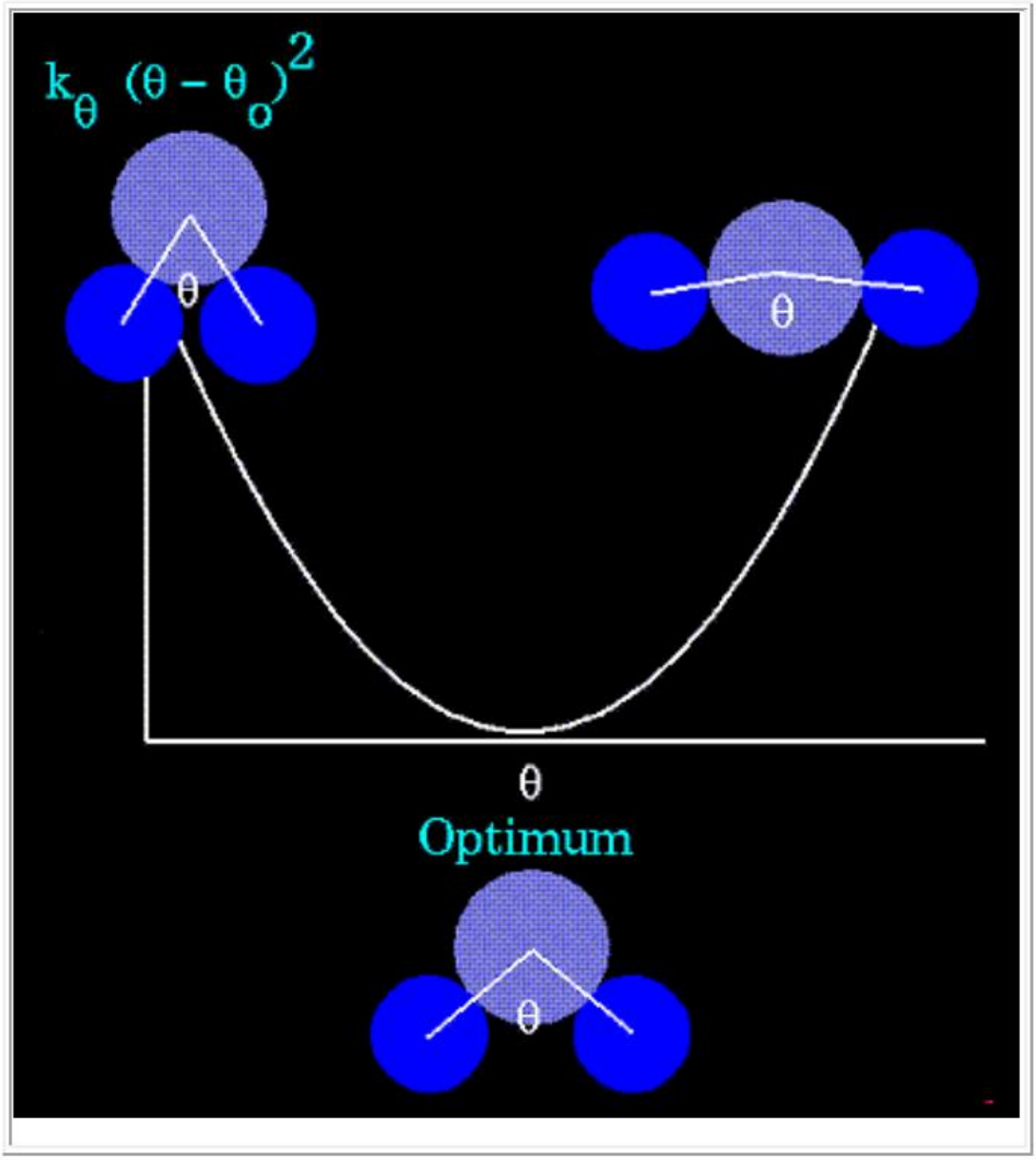
Notice that the model tends to break down as a bond is stretched toward the point of dissociation.

Bending Energy

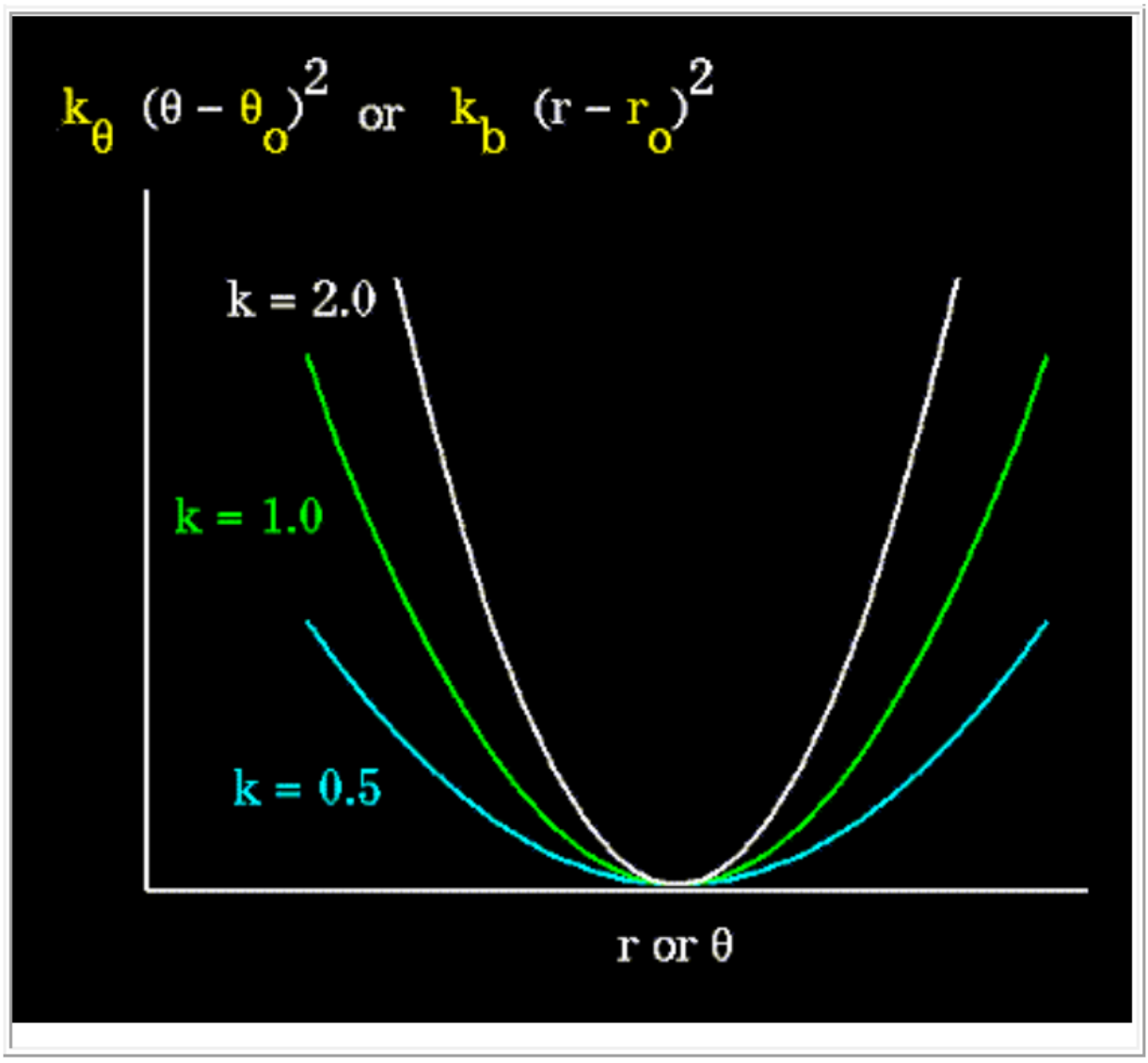
$$E = \sum_{angles} k_\theta \, (\theta - \theta_o)^2$$



The bending energy equation is also based on Hooke's law. The "k*theta*" parameter controls the stiffness of the angle spring, while "theta*o*" defines its equilibrium angle. This equation estimates the energy associated with vibration about the equilibrium bond angle:

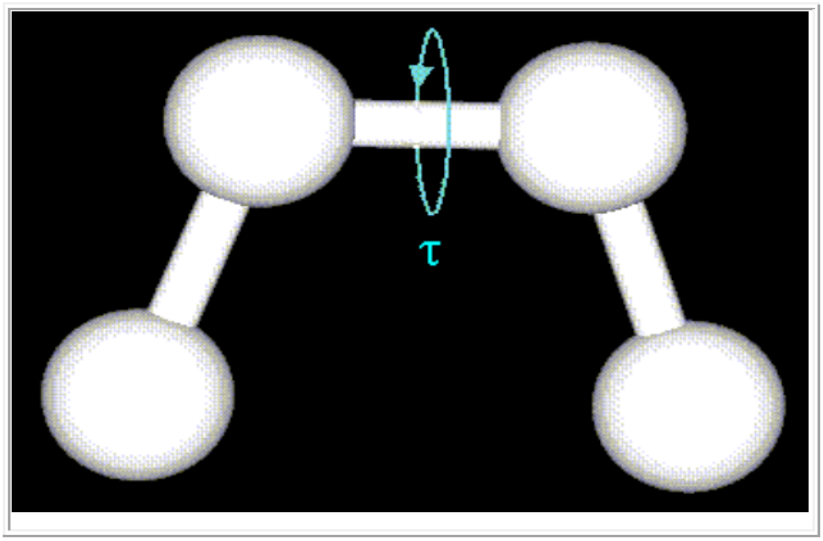


Unique parameters for angle bending are assigned to each bonded triplet of atoms based on their types (e.g. C-C-C, C-O-C, C-C-H, etc.). The effect of the "k*b*" and "k*theta*" parameters is to broaden or steepen the slope of the parabola. The larger the value of "k", the more energy is required to deform an angle (or bond) from its equilibrium value. Shallow potentials are achieved for "k" values between 0.0 and 1.0. The Hookeian potential is shown in the following plot for three values of "k":

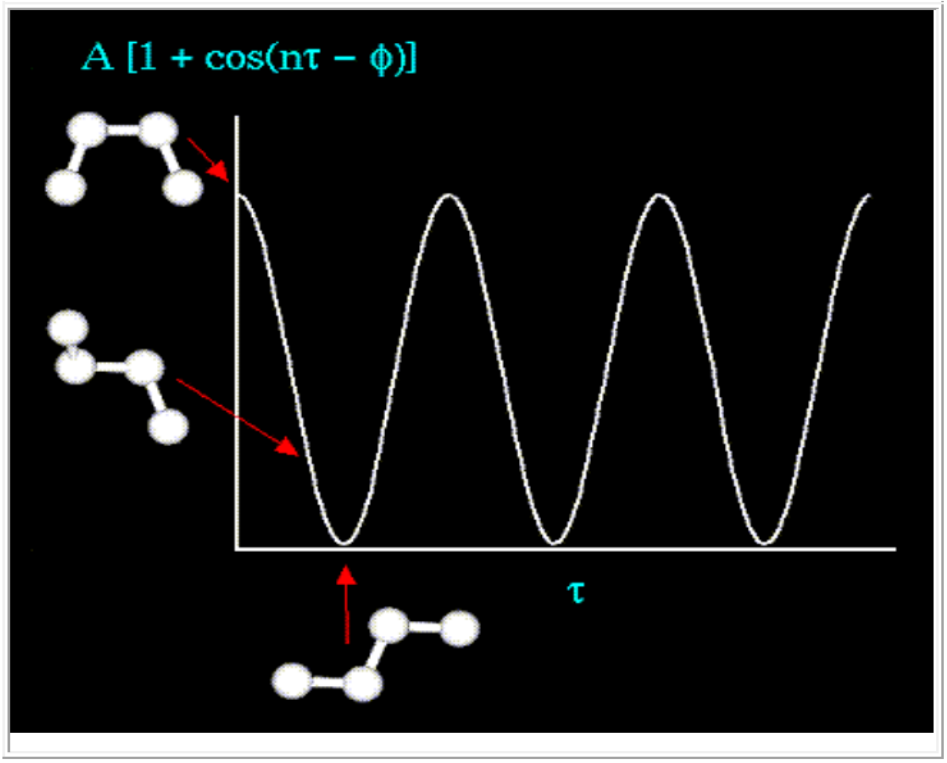


Torsion Energy

$$E = \sum_{\text{torsions}} A\,[1 + \cos(n\tau - \phi)]$$



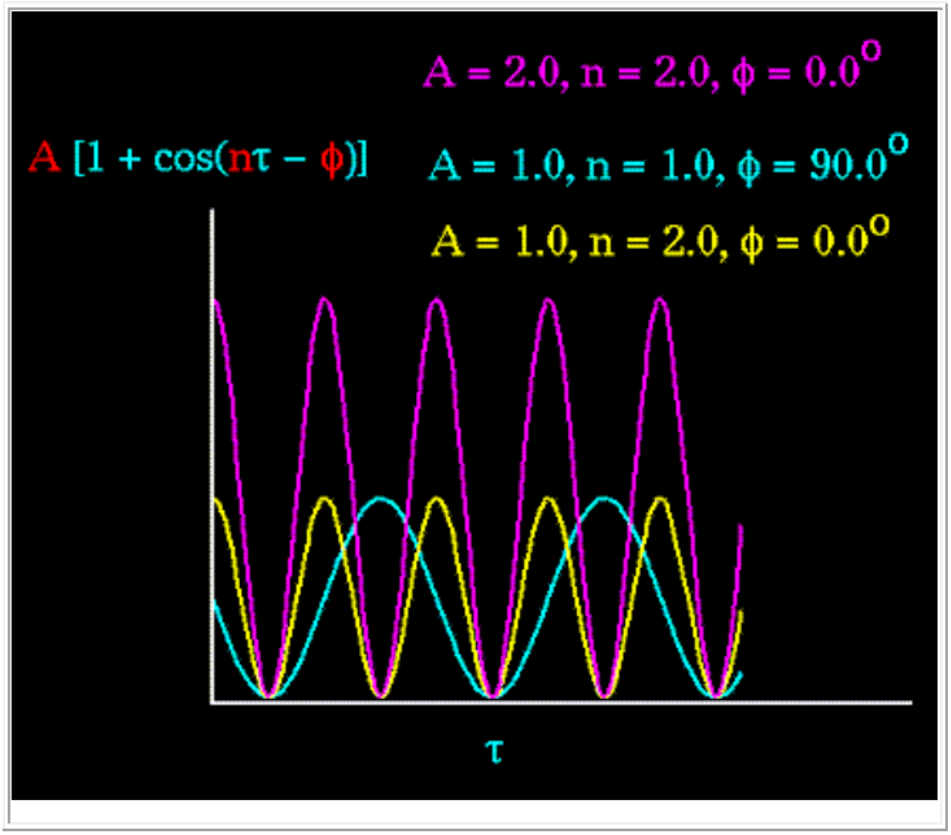
The torsion energy is modeled by a simple periodic function, as can be seen in the following plot:



The torsion energy in molecular mechanics is primarily used to correct the remaining energy terms rather than to represent a physical process. The torsional energy represents the amount of energy that must be added to or subtracted from the Stretching Energy + Bending Energy + Non-Bonded Interaction Energy terms to make the total energy agree with experiment or

rigorous quantum mechanical calculation for a model dihedral angle (ethane, for example might be used a a model for any H-C-C-H bond).

The "A" parameter controls the amplitude of the curve, the n parameter controls its periodicity, and "phi" shifts the entire curve along the rotation angle axis (tau). The parameters are determined from curve fitting. Unique parameters for torsional rotation are assigned to each bonded quartet of atoms based on their types (e.g. C-C-C-C, C-O-C-N, H-C-C-H, etc.). Torsion potentials with three combinations of "A", "n", and "phi" are shown in the following plot:



Notice that "n" reflects the type symmetry in the dihedral angle. A CH3-CH3 bond, for example, ought to repeat its energy every 120 degrees. The *cis* conformation of a dihedral angle is assumed to be the zero torsional angle by convention. The parameter phi can be used to synchronize the torsional potential to the initial rotameric state of the molecule whose energy is being computed.

Cross terms

The presence of cross terms in a forcefield reflects coupling between the internal coordinates. For example, as a bond angle is decreased it is found that the adjacent bonds stretch to reduce the interaction between the 1,3 atoms, as illustrated in Figure.

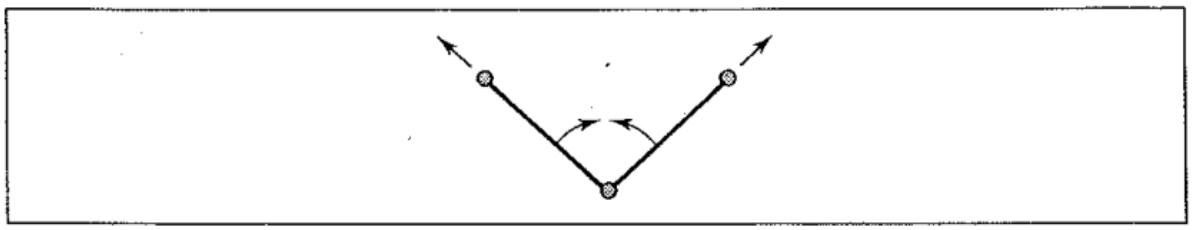
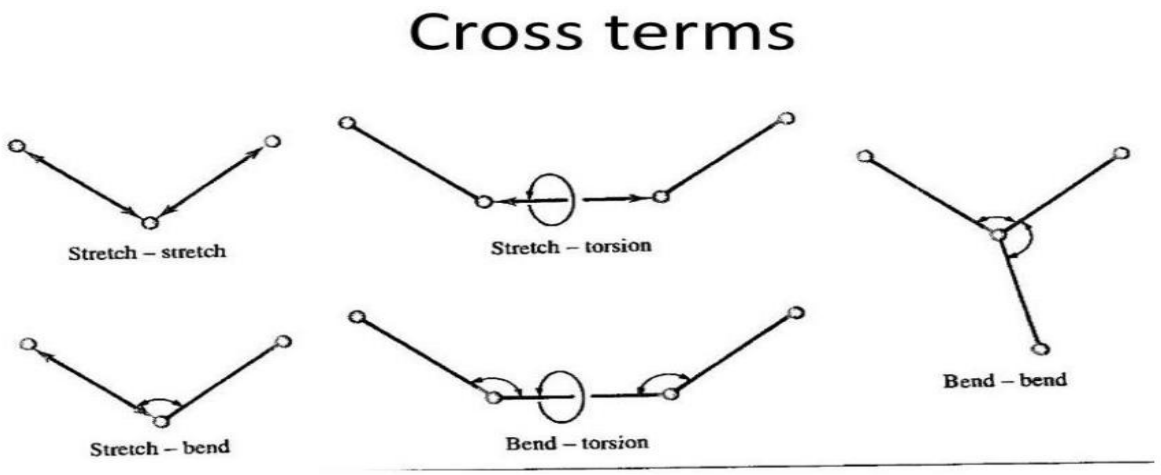


Fig. 4.12: Coupling between the stretching of the bonds as an angle closes.

One should in principle include cross terms between all contributions to a force field. However, only a few cross terms are generally found to be necessary in order to reproduce structural properties accurately; more may be needed to reproduce other properties such as vibrational frequencies, which are more sensitive to the presence of such terms. In general, any interactions involving motions that are far apart in a molecule can usually be set to zero. Most cross terms are functions of two internal coordinates, such as stretch-stretch, stretch-bend and stretch-torsion terms, but cross terms involving more than two internal coordinates such as the bend-bend-torsion have also been used.



Various functional forms are possible for the cross terms. For example, the stretch-stretch cross term between two bonds 1 and 2 can be modelled as:

$$v(l_1, l_2) = \frac{k_{l_1,l_2}}{2}[(l_1 - l_{1,0})(l_2 - l_{2,0})] \quad (4.13)$$

The stretching of the two bonds adjoining an angle could be modelled using an equation of the following form (as in MM2, MM3 and MM4):

$$v(l_1, l_2, \theta) = \frac{k_{l_1,l_2,\theta}}{2}[(l_1 - l_{1,0}) + (l_2 - l_{2,0})](\theta - \theta_0) \quad (4.14)$$

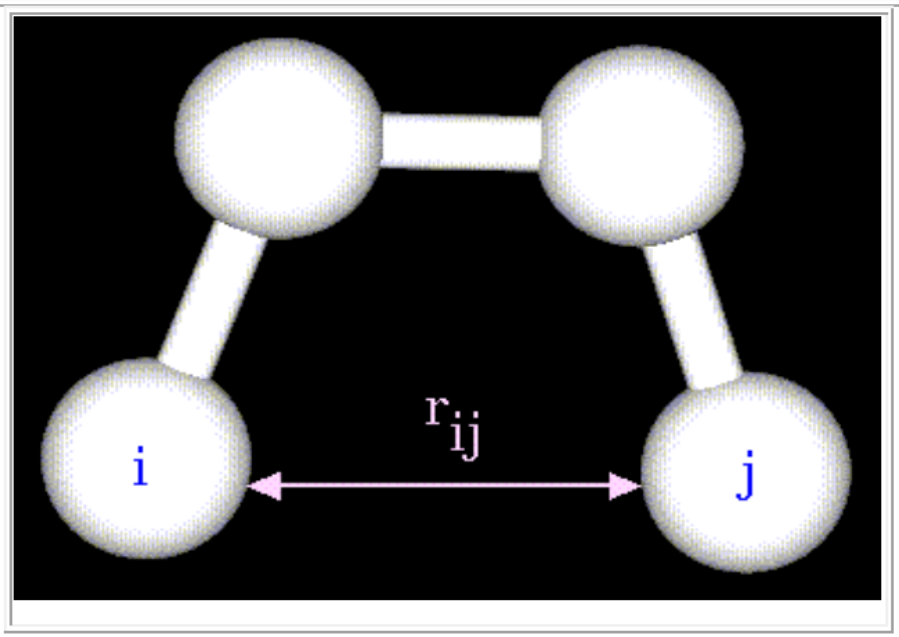
Non-Bonded Energy

Independent molecules and atoms interact through non-bonded forces, which also play an important role in determining the structure of individual molecular species. The non-bonded interactions do not depend upon a specific bonding relationship between atoms. They are 'through-space' interactions and are usually modelled as a function of some inversepower of the distance. The non-bonded terms in a forcefield are usually considered in two groups, one comprising electrostatic interactions and the other van der Waals interactions.

The non-bonded energy represents the pair-wise sum of the energies of all possible interacting non-bonded atoms i and j:

$$E = \sum_i \sum_j \frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} + \sum_i \sum_j \frac{q_i q_j}{r_{ij}}$$

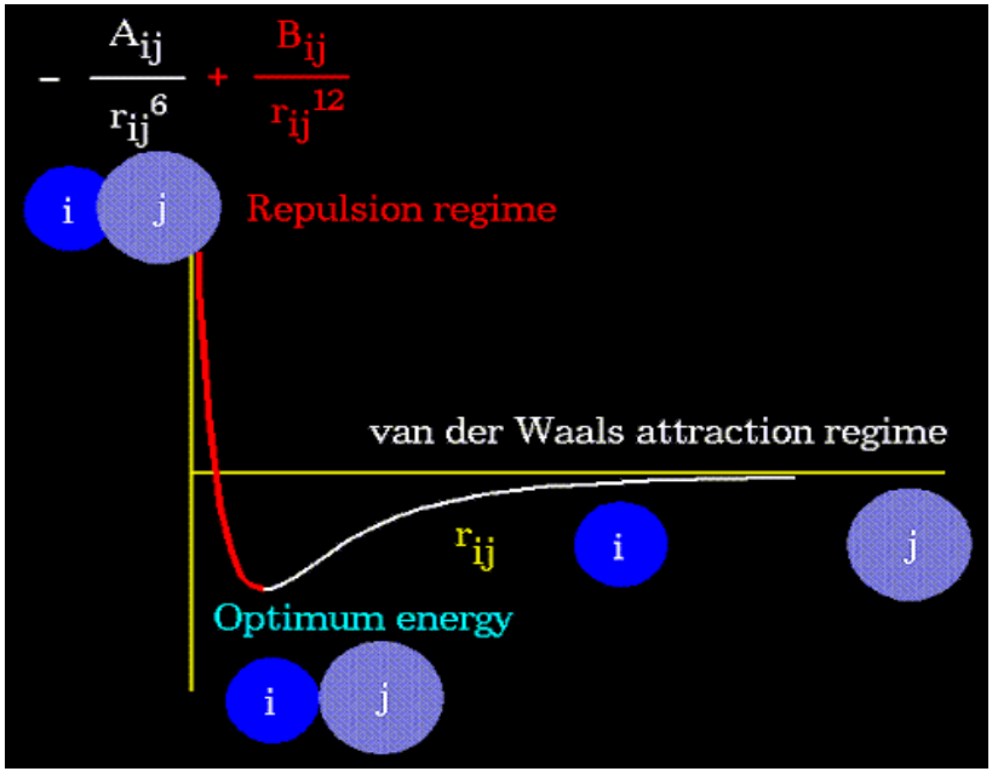
van der Waals term Electrostatic term



The non-bonded energy accounts for repulsion, van der Waals attraction, and electrostatic interactions.

Van der Waals attraction occurs at short range, and rapidly dies off as the interacting atoms move apart by a few Angstroms. Repulsion occurs when the distance between interacting atoms becomes even slightly less than the sum of their contact radii. Repulsion is modeled by an equation that is designed to rapidly blow up at close distances. The energy term that describes attraction/repulsion provides for a smooth transition between these two regimes. These effects are often modeled using a 6-12 equation, as shown in the following plot:

The "A" and "B" parameters control the depth and position (interatomic distance) of the potential energy well for a given pair of non-bonded interacting atoms (e.g. C:C, O:C, O:H, etc.). In effect, "A" determines the degree of "stickiness" of the van der Waals attraction and "B" determines the degree of "hardness" of the atoms (e.g marshmallow-like, billiard ball-like, etc.).



Vanderwaals interaction

- Dispersive interactions- long range attractive forces
- Due to instantaneous dipoles which arise due to fluctuation in electron clouds
- This can induce a dipole in neighboring atoms – giving rise to an attractive inductive effect

A simple model to explain the dispersive interaction was proposed by Drude. This model consists of 'molecules' with two charges, $+q$ and $-q$, separated by a distance r. The negative charge performs simple harmonic motion with angular frequency ω along the z axis about the stationary positive charge (Figure 4.33). If the force constant for the oscillator is k and if the mass of the oscillating charge is m, then the potential energy of an isolated Drude molecule is $\frac{1}{2}kz^2$, where z is the separation of the two charges. ω is related to the force constant by $\omega = \sqrt{k/m}$. The Schrödinger equation for a Drude molecule is:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial z^2}+\frac{1}{2}kz^2\psi = E\psi \qquad (4.59)$$

This is the Schrödinger equation for a simple harmonic oscillator. The energies of the system are given by $E_\nu = (\nu+\frac{1}{2}) \times \hbar\omega$ and the zero-point energy is $\frac{1}{2}\hbar\omega$.

Electrostatic interactions

Electrostatic interactions also arise from changes in the charge distribution of a molecule or atom caused by an external field, a process called polarisation. The primary effect of the external electric field (which in our case will be caused by neighbouring molecules) is to induce a dipole in the molecule. The magnitude of the induced dipole moment μind is proportional to the electric field E, with the constant of proportionality being the polarisability a:

$$\boldsymbol{\mu}_{\text{ind}} = \alpha \mathbf{E} \tag{4.51}$$

The energy of interaction between a dipole $\boldsymbol{\mu}_{\text{ind}}$ and an electric field **E** (the induction energy) is determined by calculating the work done in charging the field from zero to *E*, using the following integral:

$$v(\alpha, E) = -\int_0^E d\mathbf{E}\,\boldsymbol{\mu}_{\text{ind}} = -\int_0^E d\mathbf{E}\,\alpha\mathbf{E} = -\tfrac{1}{2}\alpha E^2 \tag{4.52}$$

In strong electric fields contributions to the induced dipole moment that are proportional to E^2 or E^3 can also be important, and higher-order moments such as quadrupoles can also be induced. We will not be concerned with such contributions.

The electrostatic contribution is modeled using a Coulombic potential. The electrostatic energy is a function of the charge on the non-bonded atoms, their interatomic distance, and a molecular dielectric expression that accounts for the attenuation of electrostatic interaction by the environment (e.g. solvent or the molecule itself). Often, the molecular dielectric is set to a constant value between 1.0 and 5.0. A linearly varying distance-dependent dielectric (i.e. 1/r) is sometimes used to account for the increase in environmental bulk as the separation distance between interacting atoms increases.

- Central multipole expansion
  - Electronegative elements attract electrons
  - Unequal charge distribution – fractional point charges through out the mol
  - Charges – produce the electrostatic potential
  - Charges restricted to nuclear centres – partial atomic charges

often referred to as *partial atomic charges* or *net atomic charges*. The electrostatic interaction between two molecules (or between different parts of the same molecule) is then calculated as a sum of interactions between pairs of point charges, using Coulomb's law:

$$\mathscr{V} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{4.19}$$

N_A and N_B are the numbers of point charges in the two molecules. This approach to the