Molecular mechanics

classical description of molecules

Marcus Elstner and Tomáš Kubař

April 29, 2016

Chemical bond

- quantum effect
- solution of the SR – numerically expensive (only small molecules can be treated)
- approximations necessary

Born–Oppenheimer approximation

- electrons adjust to nuclei instantaneously
- only ground state populate

classical treatment of nuclei

- 'field' for the moving electrons
- neglect of nuclear quantum effects (tunneling, ZPV)

Chemical bond

- difficult – delocalization of electron (as in metal)
- often – well localized bonding orbitals (organic molecules)
- idea – similar bonds have similar strength and properties
  e.g. similar C–H σ-orbitals $\rightarrow$ all C–H bonds are 'similar'

How to model it?

Chemical bond

- harmonic spring – the simplest possible function

$$\begin{aligned} E(x) &= \frac{1}{2}k(x-x_0)^2 \\ F(x) &= -\frac{\partial E(x)}{\partial x} = -k(x-x_0) \end{aligned}$$

- 2 parameters k and x_0 with defined meaning
  – can be obtained from spectroscopy
- can work very well for diatomics (H_2, O_2 or N_2)

Diatomics – parametrization

- 1D system – equiv. to 1 particle with m + spring k
- Newton's second law:

$$F = ma = m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -k(x - x_0)$$

- solution:

$$x(t) - x_0 = c_1 \cdot \sin\left[\sqrt{\frac{k}{m}}\,t\right] + c_2 \cdot \cos\left[\sqrt{\frac{k}{m}}\,t\right]$$

Diatomics – parametrization

- 1D system – equiv. to 1 particle with m + spring k
- Newton's second law:

$$F = ma = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -k(x - x_0)$$

- solution:

$$x(t) - x_0 = c_1 \cdot \sin\left[\sqrt{\frac{k}{m}}\,t\right] + c_2 \cdot \cos\left[\sqrt{\frac{k}{m}}\,t\right]$$

- eqn for harmonic motion

$$x(t) = c \cdot \sin(\omega t)$$

- relation of harmonic frequency with k and m:

$$\omega = \sqrt{\frac{k}{m}}$$

Diatomics – parametrization

- 2nd derivative of energy equals spring constant

$$\frac{\mathrm{d}^2 E(x)}{\mathrm{d}x^2} = \frac{\mathrm{d}^2 \frac{1}{2} k x^2}{\mathrm{d}x^2} = k$$

- generally: 2nd derivative of energy determines the (harmonic) frequency of motion of atoms in molecules – vibrations
- parametrization of the simple force field is necessary
  - based on experimental data
    (equilibrium distances and vibrational frequencies)
  - or get the data from quantum chemistry

Parametrization – general

- let us put springs in place of all covalent bonds
- do we need to parametrize k and x_0 for each bond separately? (this would be uncool. . .)
- rather, use just several set of k and x_0,
  one for each 'type' of bonding situation
- the concept of type is crucial

Concept of (atom, bond. . .) type

- evidence for the ‘unification’ of force-field parameters?

Spectroscopy

- every C–H bond: length 1.06–1.11 Å,
  frequency ca. 3100 cm^{-1}, in any molecular environment

Thermochemistry

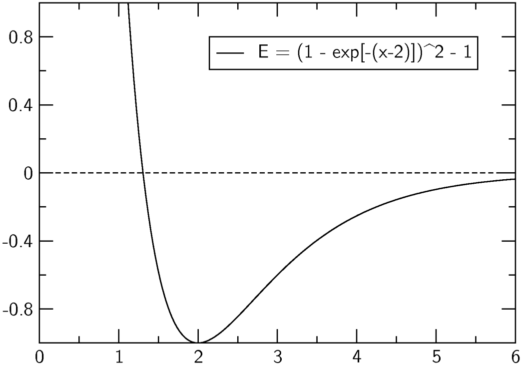
- heat of formation – roughly additive:

$$\mathrm{CH_4} \cong 4\ \mathrm{C{-}H}$$
$$\mathrm{C_2H_6} \cong 6\ \mathrm{C{-}H} + \mathrm{C{-}C}$$

- connection to well depth in Morse’s potential

Concept of (atom, bond. . .) type

Morse's potential



Concept of (atom, bond. . .) type

Bonds between atoms to be modeled by universal functions

How to identify the atom types? – chemical ideas

Concept of (atom, bond. . .) type

Bonds between atoms to be modeled by universal functions
How to identify the atom types? – chemical ideas
i) hybridization

- difference between sp^3 carbon (4 bonds) and sp^2 C (3 bonds) – these will be the different atom types
- different functions for bonds of types C–C, C=C and C≡C
- determine the parameters (k, x_0) with some selected molecules, typical for the binding situation
- example: use C_2H_6, C_2H_4, C_2H_2 and benzene for k, x_0

Concept of (atom, bond. . .) type

Bonds between atoms to be modeled by universal functions
How to identify the atom types? – chemical ideas
ii) polarity

- an atom bonded to electronegative atom – electron deficient
  → affects its bonding to other atoms
- example: O=CH–C. . .
  the C–C bond is affected and needs to be parametrized
  differently from apolar C–C
  → an atom type for carbonyl C introduced

Biomolecular force fields
– usually 20 types for C, 10 for N and 5 for O and H

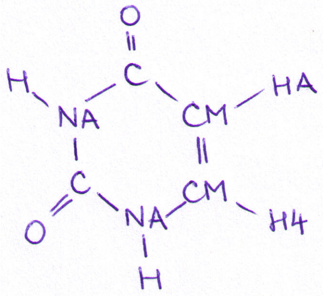
Concept of (atom, bond. . .) type

AMBER force field for biomolecules:

C - sp2 C carbonyl group
CA - sp2 C pure aromatic (benzene)
CB - sp2 aromatic C, 5&6 membered ring junction
CC - sp2 aromatic C, 5 memb. ring HIS
CK - sp2 C 5 memb.ring in purines
CM - sp2 C pyrimidines in pos. 5 & 6
CN - sp2 C aromatic 5&6 memb.ring junct.(TRP)
CQ - sp2 C in 5 mem.ring of purines between 2 N
CR - sp2 arom as CQ but in HIS
CT - sp3 aliphatic C
CV - sp2 arom. 5 memb.ring w/1 N and 1 H (HIS)
CW - sp2 arom. 5 memb.ring w/1 N-H and 1 H (HIS)
C* - sp2 arom. 5 memb.ring w/1 subst. (TRP)

Concept of (atom, bond. . .) type

AMBER atom types in a molecule of uracil



The localization of wave function / charge

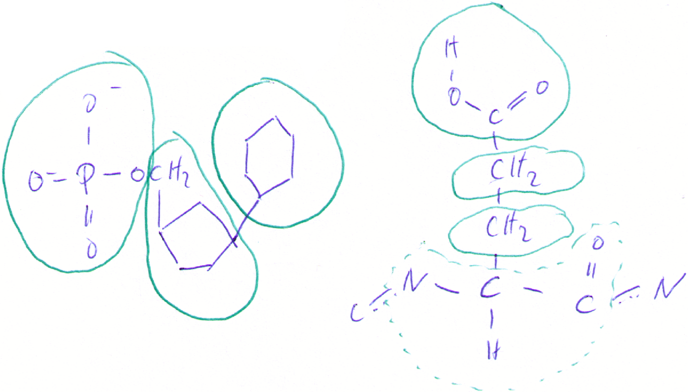
- quantum-mechanical basis of parametrized force field
- localized in a bond or in a molecular fragment
  (functional group, larger pieces of biomolecule)

Importance:

- atom types – various types for differently bound carbon
- electrostatic interactions – every atom carries a point electric charge, and groups of atoms based on chemical intuition may be considered electroneutral (or charged, $+1$ or -1)

The localization of wave function / charge

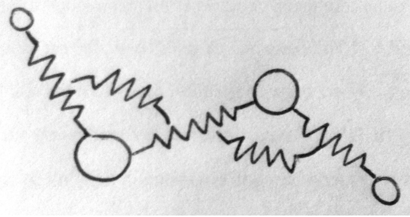
groups of atoms with integer charge in biomolecules



Interactions between atoms

Bonded

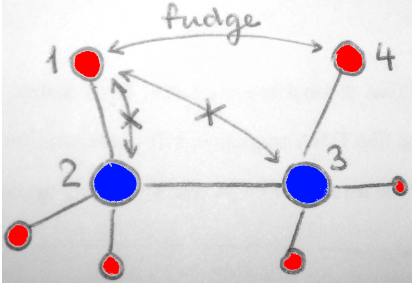
- mediated by, and resulting directly from covalent bonds
- usually described by harmonic springs between atoms, and bond and dihedral angles in a similar way
- cover all of the quantum-mechanical phenomena like exchange and correlation between two atoms using an effective potential function



Interactions between atoms

Non-bonded

- longer-range interactions
  - – charge–charge Coulomb and van der Waals (vdW)
- between molecules and distant parts of one molecule
- excluded between bonded atoms, up to 4th neighbor
- 1–4 interactions – scaled down (‘fudge’ factor)



Coulomb parameters

3 contributions (quantum-chemical picture)

- nucleus–nucleus

$$\frac{1}{2}\sum_{ij}\frac{Z_i \cdot Z_j}{R_{ij}}$$

- nucleus–electron with approx. point charges $\rho(r) \rightarrow \sum_j Q_j$

$$-\sum_i \int \frac{Z_i \cdot \rho(r)}{|R_i - r|}\mathrm{d}r \rightarrow -\sum_{ij}\frac{Z_i \cdot Q_j}{R_{ij}}$$

- (Hartree) electron–electron with point charges

$$\frac{1}{2}\iint \frac{\rho(r)\cdot\rho(r')}{|r - r'|} \rightarrow \frac{1}{2}\sum_{ij}\frac{Q_i \cdot Q_j}{R_{ij}}$$

Coulomb parameters

3 contributions (quantum-chemical picture)

- nucleus–nucleus
- nucleus–electron with approx. point charges $\rho(r) \rightarrow \sum_j Q_j$
- (Hartree) electron–electron with point charges

define an effective atomic charge as $q_i = -Q_i + Z_i$
then

$$E_{QQ} = \frac{1}{4\pi\varepsilon_0} \cdot \sum_i \sum_{j>i} \frac{q_i \cdot q_j}{R_{ij}}$$

Coulomb parameters

effective atomic charge needs to be defined for every atom
(rather than atom type – this would be too crude)

- from quantum-chemical calculations – straightforward
- do calculations of typical (bio)molecular fragments
  and get charges from there
- fragments
  – amino acid residues and peptide bonds for proteins
  – nucleobases, sugars and phosphate groups for DNA/RNA
- there are 2 issues. . .

Coulomb parameters – issue 1

atomic charges – no unique definition in quantum chemistry (Mulliken, or may it be NBO or AIM today?)

- several schemes proposed – hard to say, which is the right
- the simplest (Mulliken charges) do not work

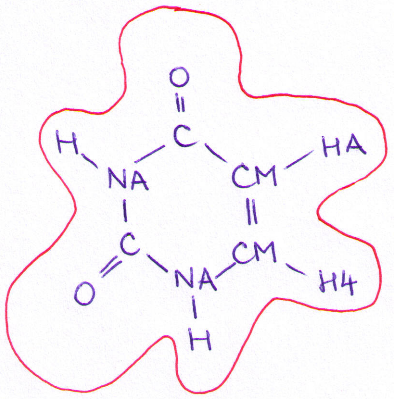
Coulomb parameters – issue 1

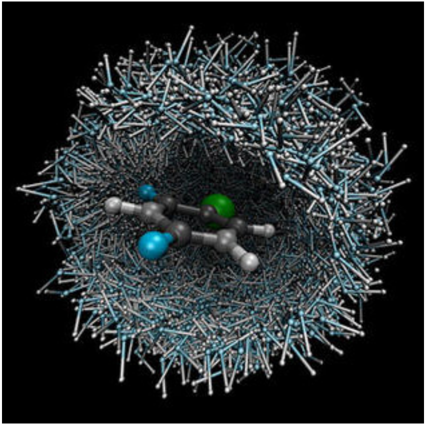
atomic charges – no unique definition in quantum chemistry (Mulliken, or may it be NBO or AIM today?)

- several schemes proposed – hard to say, which is the right
- the simplest (Mulliken charges) do not work
- popular – potential-derived charges:
  1. calculate the electron density in the molecule
  2. get electrostatic potential at surface of the molecule
  3. fit point electric charges on atoms to reproduce the ESP

Coulomb parameters – issue 1

surface of the uracil molecule





Coulomb parameters – issue 2

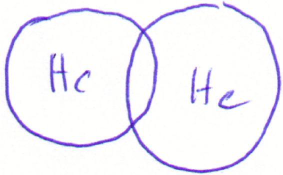
atom charges are calculated for an isolated molecule (gas phase)

- electrostatic situation in aqueous/polar solution differs
  – molecules are more polar
- example: H_2O molecule in the gas phase: 1.8 D
  in water: 2.4 D
- a good idea – to set the charges to larger values
- popular trick – use a small basis sets in QCh calculations
  – then, charges and dipole moments are overestimated
  – good for force fields w/o polarizability and charge transfer
- another possibility – polarizable force field
  atomic polarizability α_i is assigned to every atom i
  external field induces atomic dipole $\overrightarrow{\mu_i} = \overleftrightarrow{\alpha_i} \cdot \overrightarrow{E}$

van der Waals interaction

Pauli repulsion

- electrons with the same spin avoid spatial overlap
- purely quantum-mechanical phenomenon (exchange effect)
- example: interaction of neutral closed-shell systems – He_2



- electron densities start to overlap → energy grows steeply
- modeling:

$$E_{ex} = \exp\left[a - b \cdot R_{ij}\right]$$

$$E_{ex} = \left(\frac{\sigma}{R_{ij}}\right)^{12}$$

van der Waals interaction

dispersion due to correlation

- correlation – between electrons irrespective of spin,
  retained on longer distances
- fluctuation of density (even at zero point)
  → instantaneous dipole
- another dipole is induced on the other atom/molecule
- interaction between instantaneous dipole and induced dipole
- the orientation of dipoles is correlated – attractive interaction

note: dispersion is difficult to capture with quantum chemistry

van der Waals interaction

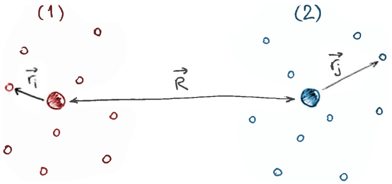
dispersion due to correlation

www.petrhejna.cz

news.softpedia.com

van der Waals interaction

dispersion – quantification (1)



i, j are electrons with charges e; $\vec{R}||\vec{z}$; $\vec{r}_i = (x_i, y_i, z_i)$; $X_1 = \sum_i x_i$
interaction of charge distributions (leading Taylor term):

$$V = \frac{1}{R^3} \sum_{i,j} e_i e_j (x_i x_j + y_i y_j - 2 z_i z_j) = \frac{e^2}{R^3} (X_1 X_2 + Y_1 Y_2 - 2 Z_1 Z_2)$$

perturbation theory – interaction of dipoles is the perturbation

van der Waals interaction

dispersion – quantification (2)

no permanent dipole → first-order energy vanishes:

$$E^{(1)} = \left\langle \Psi^{(0)} \middle| V \middle| \Psi^{(0)} \right\rangle = 0$$

second-order energy w/ unperturbed states $\left|\Psi_m^{(0)}\right\rangle$ & energies $E_m^{(0)}$:

$$E^{(2)} = \sum_{m \neq n} \frac{\left| \left\langle \Psi_m^{(0)} \middle| V \middle| \Psi_n^{(0)} \right\rangle \right|^2}{E_n^{(0)} - E_m^{(0)}}$$

recall $V = \frac{e^2}{R^3}(X_1 X_2 + Y_1 Y_2 - 2 Z_1 Z_2)$

and write $X_{m0} = \left\langle \Psi_m^{(0)} \middle| \frac{eX}{R^3} \middle| \Psi_0^{(0)} \right\rangle$

van der Waals interaction

dispersion – quantification (3)

$$E^{(2)} = \frac{e^4}{R^6} \sum_{m \neq n} \frac{\left| X^1_{m0} X^2_{n0} + Y^1_{m0} Y^2_{n0} - 2 Z^1_{m0} Z^2_{n0} \right|^2}{E^{(0)}_{10} - E^{(0)}_{1m} + E^{(0)}_{20} - E^{(0)}_{2n}}$$

R^{-6} arises from the matrix element being squared

simplification – averaging over all orientations of clusters 1 and 2, and also considering spherically symmetric charge distributions:

$$E^{(2)} = \frac{2}{3} \frac{e^4}{R^6} \sum_{m \neq n} \frac{\left| R^1_{m0} \right|^2 \left| R^2_{n0} \right|^2}{E^{(0)}_{10} - E^{(0)}_{1m} + E^{(0)}_{20} - E^{(0)}_{2n}}$$

van der Waals interaction

dispersion – quantification (4)
consider polarizability $\alpha(0)$ in a static field:

$$\alpha(0) = \frac{2}{3} e^2 \sum_j \frac{\left|R_{j0}\right|^2}{E_j - E_0} \approx \frac{2}{3} \frac{e^2}{I} \sum_j \left|R_{j0}\right|^2$$

with ionization energy I approximating excitation energy $E_j - E_0$

$$\begin{aligned} E^{(2)} &\approx -\frac{2}{3} \frac{e^4}{R^6} \frac{1}{I_1 + I_2} \sum_{m \neq n} \left|R^1_{m0}\right|^2 \left|R^2_{n0}\right|^2 \\ &\approx -\frac{2}{3} \frac{I_1 I_2}{I_1 + I_2} \frac{\alpha_1(0)\, \alpha_2(0)}{R^6} = -\frac{C_6}{R^6} \end{aligned}$$

Note – higher-order terms have been neglected

van der Waals interaction

most common function: Lennard-Jones 12-6 potential

$$V(r) = 4\varepsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right)$$

- 2 parameters – σ and ε
- minimum at $r_m = 2^{1/6}\sigma$ and well depth ε

van der Waals interaction

most common function: Lennard-Jones 12-6 potential

$$V(r) = 4\varepsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right)$$

- 2 parameters – σ and ε
- minimum at $r_m = 2^{1/6}\sigma$ and well depth ε

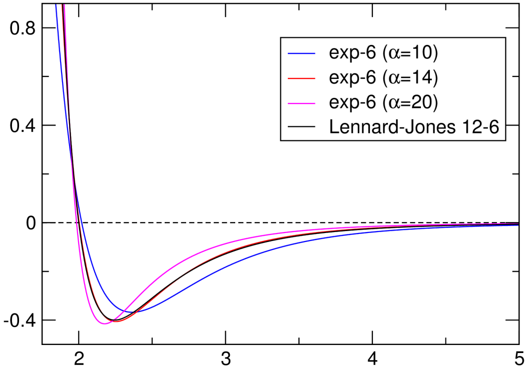
repulsive: $\exp[-r]$ sometimes better than r^{-12}
hskip 5 mm $\rightarrow$ exp-6 potential a.k.a. Buckingham potential

$$V(r) = \frac{\varepsilon}{1 - 6/\alpha}\cdot\left(\frac{6}{\alpha}\cdot\exp\left[\alpha\left(1 - \frac{r}{\sigma}\right)\right] - \left(\frac{\sigma}{r}\right)^{6}\right)$$

- α controls the steepness of the repulsive branch
- may be a better choice for phase transitions
- e.g. MM water would not freeze below 0 °C with LJ 12-6

van der Waals interaction

Lennard-Jones 12-6 potential with $\varepsilon = 0.4$ and $\sigma = 2$, and several similar exp-6 potentials with various values of α



van der Waals interaction

heteronuclear interactions A–B

– mixing rules for the parameters

$$\begin{aligned}\sigma_{\mathrm{AB}} &= \frac{\sigma_{\mathrm{AA}} + \sigma_{\mathrm{BB}}}{2} \\ \varepsilon_{\mathrm{AB}} &= \sqrt{\varepsilon_{\mathrm{AA}} \cdot \varepsilon_{\mathrm{BB}}}\end{aligned}$$

parametrization

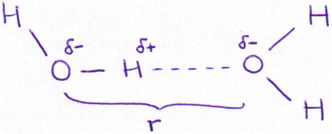
- challenging task in general
- fitting of params to experimental / quantum-chem. data
- e.g. relation to density and heat of vaporization
  – obvious in organic liquids – major interaction

Hydrogen bonding

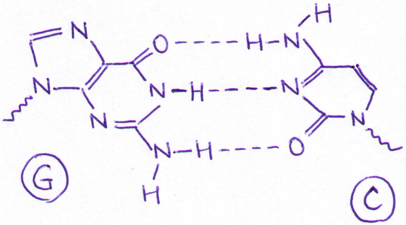
- attractive interaction between an H bonded to an electronegative atom (O, N) and another O or N
- examples: H_2O dimer, nucleobase pair
- crucial interaction in biomolecules
- interplay of several kinds of interactions
- early force fields – special potential functions for H-bonding
- typical binding energies: 20 kJ/mol
  higher for strongly polarized or even charged molecules
  or if there are several H-bonds (nucleobase pairs)

Hydrogen bonding

H_2O dimer



guanine:cytosine base pair

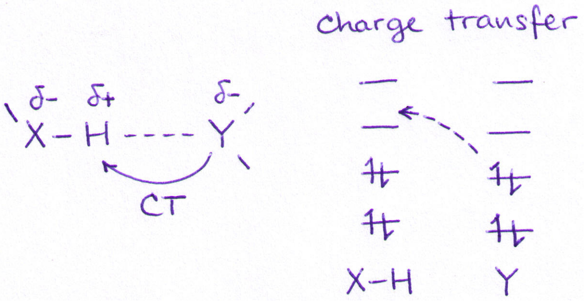


Hydrogen bonding

- Coulomb interaction is dominant
- vdW interaction
  - may become important, especially in weakly bound systems
  - crucial e.g. for angular dependence in $H_2CO\ldots H_2O$ etc.
- modern force fields – no special treatment of H-bonds
- charge transfer contribution
  - cannot be covered by force fields due to constant charges
  - may be included in other terms effectively

Hydrogen bonding

Charge transfer between the charge donor (Y) and acceptor (X–H) (or proton donor X–H and proton acceptor Y)



Parametrization of bonded interactions

Bonds

- usually represented by Morse's potential well

$$E(r) = D\left(1 - \exp\left[-\alpha(r - r_0)\right]\right)^2$$

  – not necessary, possibly inefficient, hardly used

- way to an approximated function – Taylor expansion

$$E(r) = E(r_0) + \frac{\mathrm{d}E}{\mathrm{d}r}(r_0)\cdot(r - r_0) + \frac{1}{2}\frac{\mathrm{d}^2E}{\mathrm{d}r^2}(r_0)\cdot(r - r_0)^2 + \ldots$$

Parametrization of bonded interactions

Bonds

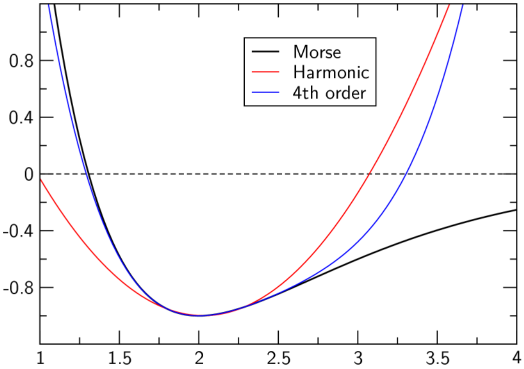
- most force fields – harmonic approximation (Taylor expansion cut after 2nd-order term)

$$E(r) = \frac{1}{2} k (r - r_0)^2$$

- approximation of the true potential with a quadratic function
  – works in a narrow interval of distances
  – sufficient for most applications as vibrations are within the interval
- if bonds are to be created or broken (chemistry)
  another solution has to be sought
  → probably leave molecular mechanics ☹

Parametrization of bonded interactions

Comparison of the harmonic and quartic functions with Morse's



- for accurate vibration frequencies – 4th-order terms can be important to describe the curvature

Parametrization of bonded interactions

Angles

- harmonic approximation for the angle deformation

$$E_{\text{bend}}(\vartheta) = \frac{1}{2} k_\vartheta (\vartheta - \vartheta_0)^2$$

- two parameters needed – equilibrium angle and force constant
- from experiment (vib-rot spectra) or quantum chemistry

Parametrization of bonded interactions

Dihedral angles

- describe the rotation around covalent bonds
- defined by 4 atoms
- potential energy – periodic function of the dihedral angle:

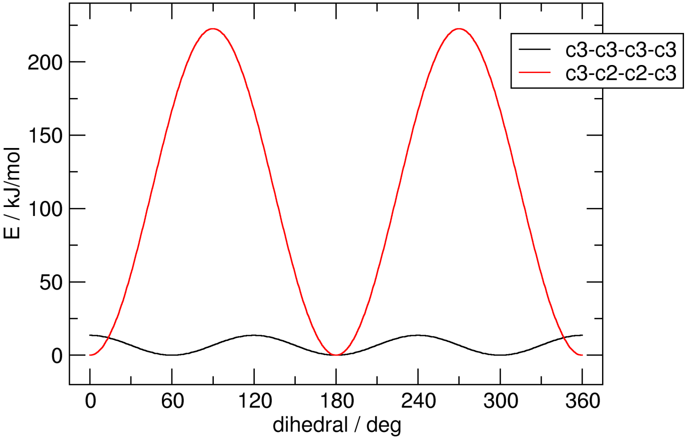
$$E(\omega) = \sum_{n=1,2,3,4,6} V_n \cos\left[n \cdot \omega - \gamma_n\right]$$

- V_n – amplitude (barrier), n – periodicity, γ – phase shift
- Ryckaert-Bellemans function – often used for alkanes:

$$E(\Omega) = \sum_{n=0}^{5} C_n \cos^n \Omega \qquad (\Omega = \omega - 180^\circ)$$

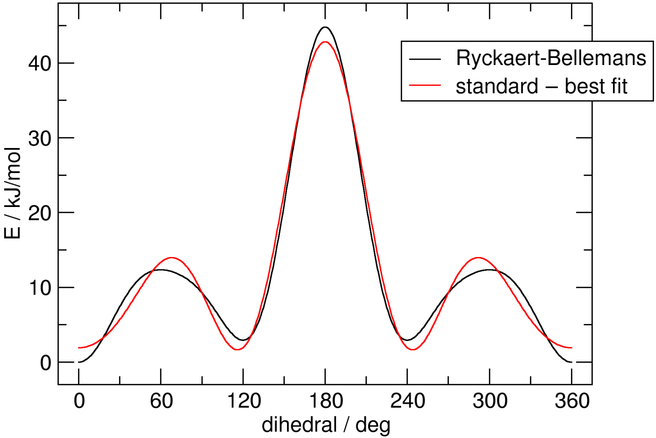
Parametrization of bonded interactions

Dihedral angles – example: C–C single and C=C double bonds



Parametrization of bonded interactions

Dihedral angles – comparison of RB dihedral with standard form



The complete equation

Adding up all contributions, the total energy of a typical biomolecular force field reads

$$E(R^N) =$$

$$\begin{aligned} = \quad & \frac{1}{2}\sum_i k_i (r_i - r_i^0)^2 + \frac{1}{2}\sum_j k_j^\vartheta (\vartheta_j - \vartheta_j^0)^2 + \frac{1}{2}\sum_n V_n \cdot \cos\left[n\omega - \gamma_n\right] \\ + \quad & \sum_i^N \sum_{j=i+1}^N \left\{ 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right) + \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \right\} \end{aligned}$$

Forces

force acting on atom i

– derivative of energy with respect to the coordinates of atom i

$$\begin{aligned} \vec{F}_i &= -\nabla_i V \\ F_i^x &= -\frac{\partial V}{\partial x_i} \end{aligned}$$

derivatives of all of the terms in the force field
are obtained in an analytical form easily

Forces

most terms depend on distances of atoms

$$V = V(r_{ij}) + V(r_{ik}) + V(r_{il}) + \dots$$

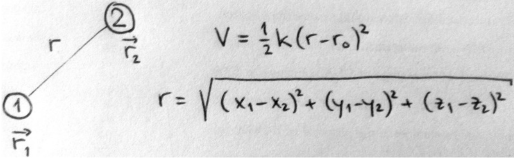
and thus the force

$$\begin{aligned} F_i^x &= -\frac{\partial V}{\partial x_i} = \\ &= -\frac{\partial V(r_{ij})}{\partial x_i} - \frac{\partial V(r_{ik})}{\partial x_i} - \frac{\partial V(r_{il})}{\partial x_i} - \dots = \\ &= -\frac{\partial V(r_{ij})}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_i} - \frac{\partial V(r_{ik})}{\partial r_{ik}} \frac{\partial r_{ik}}{\partial x_i} - \frac{\partial V(r_{il})}{\partial r_{il}} \frac{\partial r_{il}}{\partial x_i} - \dots \end{aligned}$$

(more complex derivatives for angle/dihedral contributions)

$$\frac{\partial V(\sphericalangle ijk)}{\partial \sphericalangle ijk} \frac{\partial \sphericalangle ijk}{\partial x_i}, \quad \frac{\partial V(\sphericalangle ijkl)}{\partial \sphericalangle ijkl} \frac{\partial \sphericalangle ijkl}{\partial x_i}$$

Forces – example



$$V = \frac{1}{2}k(r-r_0)^2$$

$$r = \sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$$

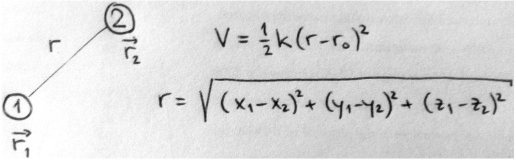
$$V(r_{12}) = \frac{1}{2}k(r_{12}-r_0)^2$$

$$\frac{\partial V(r_{12})}{r_{12}} = k(r_{12}-r_0)$$

$$r_{12} = \sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$$

$$\frac{\partial r_{12}}{\partial x_1} = \frac{1}{2r_{12}} \cdot \frac{\partial (x_1-x_2)^2}{\partial x_1} = \frac{1}{2r_{12}} \cdot 2(x_1-x_2) = \frac{x_1-x_2}{r_{12}}$$

Forces – example



$$V = \frac{1}{2}k(r-r_0)^2$$

$$r = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2 + (z_1-z_2)^2}$$

$$\frac{\partial V}{\partial x_1} = \frac{\partial V(r_{12})}{r_{12}} \frac{\partial r_{12}}{\partial x_1} = k(r_{12} - r_0) \cdot \frac{x_1 - x_2}{r_{12}}$$

$$\vec{F}_1 = -\nabla_1 V = -k(r_{12} - r_0) \cdot \frac{\vec{r}_{12}}{r_{12}} \quad \text{with } \vec{r}_{12} = \vec{r}_1 - \vec{r}_2$$