

# An Introduction to Force Fields

Prof. Michael Shirts

University of Colorado Boulder

7th i-CoMSE Workshop: Molecular Dynamics

Boise State University, July 8, 2024

Licence: CC-BY 4.0



# What is a Force Field?

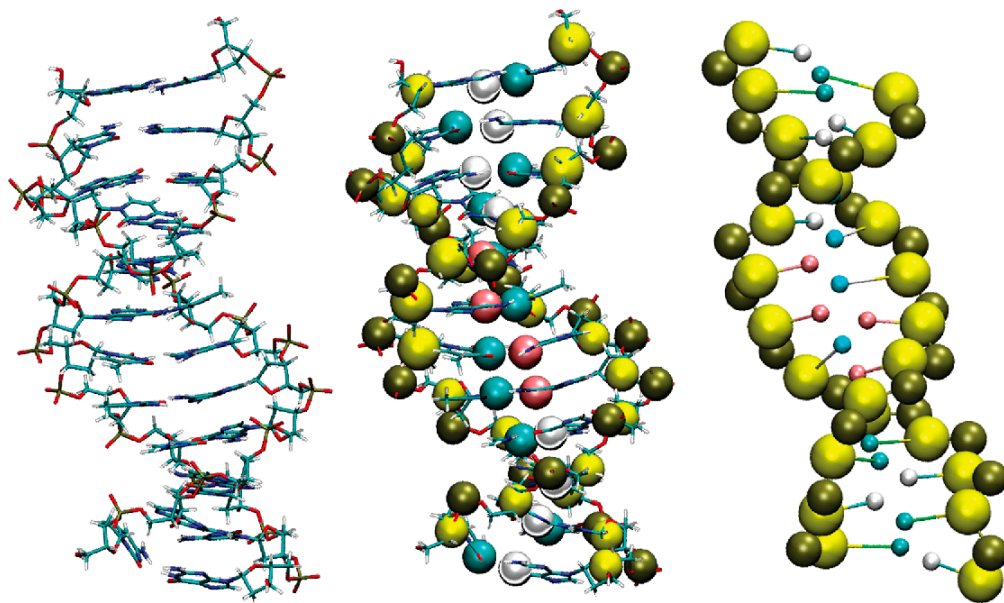
$$U(x_1, x_2, x_3, \dots, x_n)$$

- **Any way** to assign an energy to the coordinates of particles in the system
- You can compute the force and use for MD as long as the system has continuous derivatives.
- Pick a  $U(x)$  **that is good enough for your purposes**.
- If you define a  $U(x)$ , you also define an  $F(x)$ , which assigns a vector (the force) to each configuration in space.
  - Thus, it is a **vector field**, where the vector is the force!

$$F(\vec{x}) = - \left[ \frac{\partial U(\vec{x})}{\partial x_1}, \frac{\partial U(\vec{x})}{\partial x_2}, \dots, \frac{\partial U(\vec{x})}{\partial x_n} \right]$$

# What are we trying to simulate?

- We have some **model** of our molecular system.
- We have the **hypothesis** that the physical behavior only depends on **some** of the physics of the molecule
- We test that hypothesis with a simulation
- This model can be atomistic, quantum mechanical, or coarse grained.
  - In CG, each particle isn't an atom, but a methane particle, or an amino acid in a protein

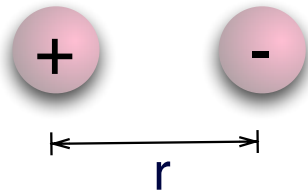


# We break up the potential energy into classical terms, and fit the resulting

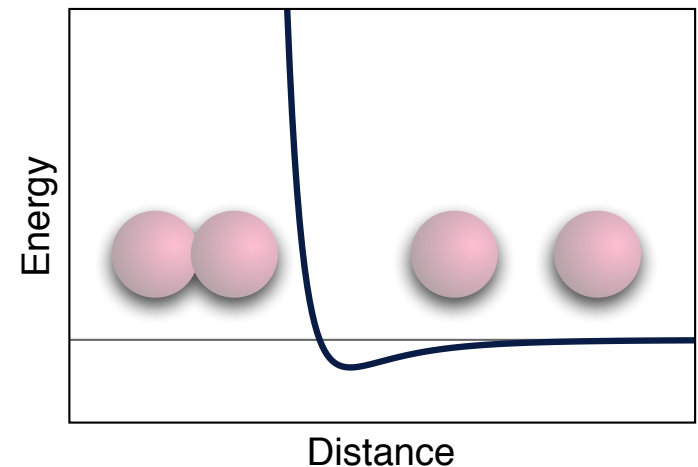
Most common nonbonded force fields are **pair potentials**

$$V(r_1, r_2, \dots, r_n) = \sum_{i < j} V_{ij}(r_{ij})$$

$$U(\mathbf{x}) = \sum_{\text{all pairs}} \frac{q_i q_j}{r}$$



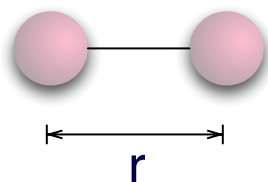
$$+ \sum_{\text{all pairs}} \frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6}$$



# "Bonded Terms"

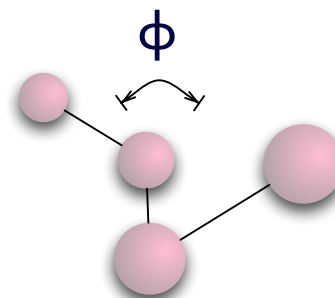
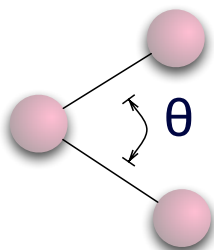
- We replace the electrons in orbitals with simple functional forms that are  $10^6$  times cheaper (but WAY more approximate).

$$+ \sum_{\text{all bonds}} \frac{1}{2} K_b (r - r_0)^2$$



$$+ \sum_{\text{all torsions}} \sum_n K_\phi [1 - \cos(n\phi + \phi_0)]$$

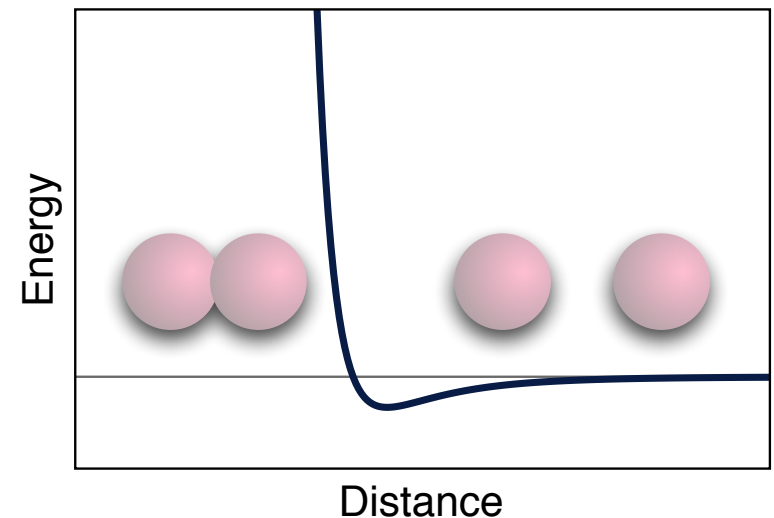
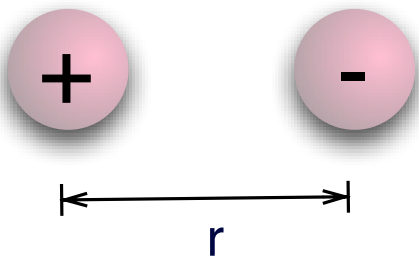
$$+ \sum_{\text{all angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2$$



# "Nonbonded terms"

- We replace the electrostatic / dipole / induced dipole interactions with physically motivated (but not exact) functional forms.

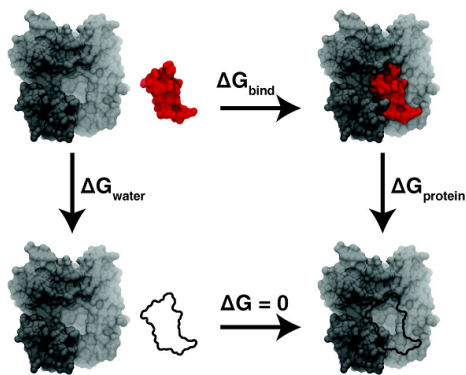
$$U(\mathbf{x}) = \sum_{\text{all pairs}} \frac{q_i q_j}{r} + \sum_{\text{all pairs}} \frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6}$$



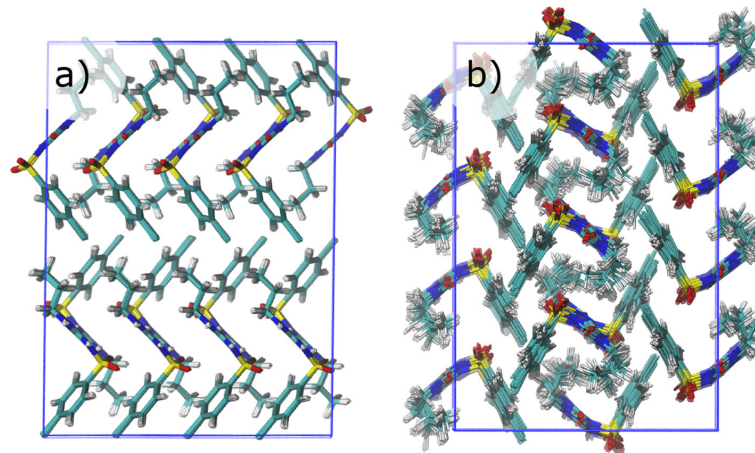
# Small molecule force fields are widely-used tools in computational biophysics and soft

- Surprisingly, all-atom fixed charge force fields strike an appropriate balance between chemical accuracy and computation efficiency for many systems of interest.

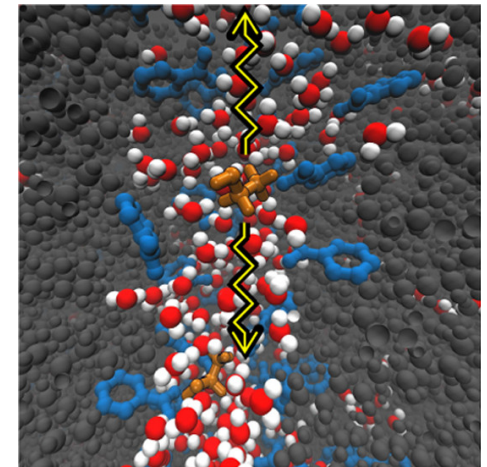
Protein-Ligand Binding



Drug molecule screening

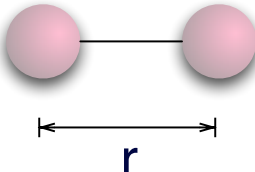
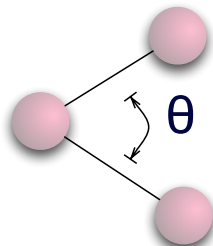


Molecular transport



# "Bonded Terms"

- We replace the energies of electrons jiggling in molecular orbitals with simple functional forms that are  $10^6$  times cheaper (but WAY more approximate).

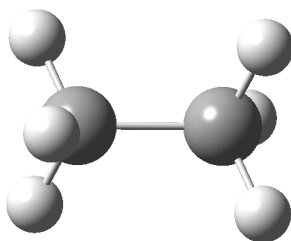
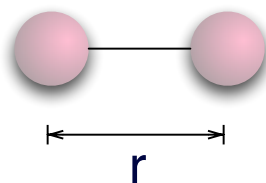
$$+ \sum_{\text{all bonds}} \frac{1}{2} K_b (r - r_0)^2$$

$$+ \sum_{\text{all angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2$$


- Most common are harmonic bonds and angles
- Pretty good approximation to quantum mechanics!
- Can get the parameters directly from QM of isolated molecules, as they are relatively independent of environment



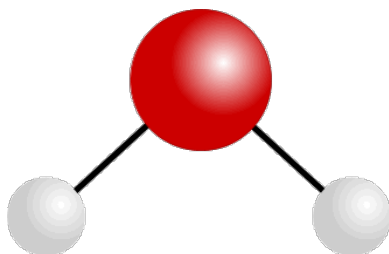
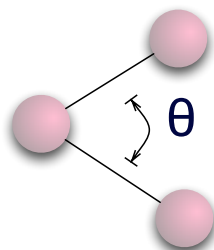
# Bonds and Angles

$$+ \sum_{\text{all bonds}} \frac{1}{2} K_b (r - r_0)^2$$



```
[ bondtypes ]
; i      j  func      b0      kb
CT      HC      1      0.10900  284512.0 ;
CA      CA      1      0.14000  392459.2 ;
CA      HA      1      0.10800  307105.6 ;
```

$$+ \sum_{\text{all angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2$$



```
[ angletypes ]
; i      j      k      func      theta0      ka
HC      CT      HC      1      107.800      276.144 ;
CA      CA      CA      1      120.000      527.184 ;
CA      CA      HA      1      120.000      292.88  ;
CA      CA      CB      1      120.000      527.184 ;
```

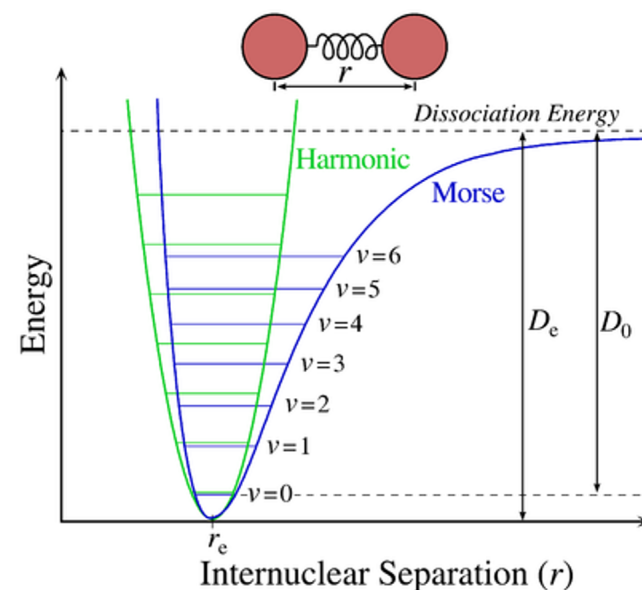
# Other Bonds and Angles Options

- Bonds

- Could use

$$U(x) = k_2(r - r_0)^2 + k_4(r - r_0)^4$$

- Morse potentials, that allow for disassociation:



- Angles:

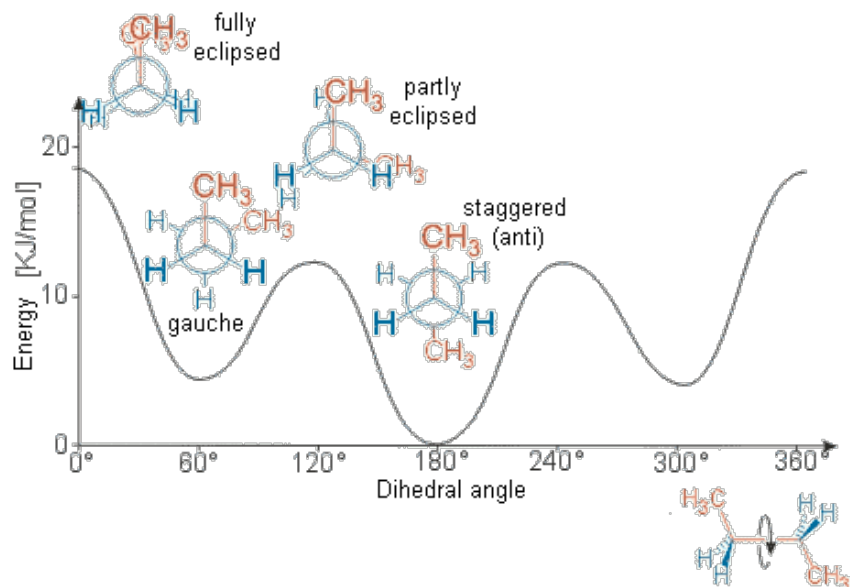
- Could add Urey-Bradley term

$$U(x) = k_\theta(\theta - \theta_0)^2 + k_{UB}(r_{ik} - r_u b)^2$$

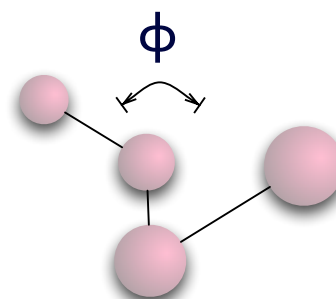
- ML potentials

# Torsions, or the junk bin of force fields

- Approximate QM with a sum of periodic trig functions



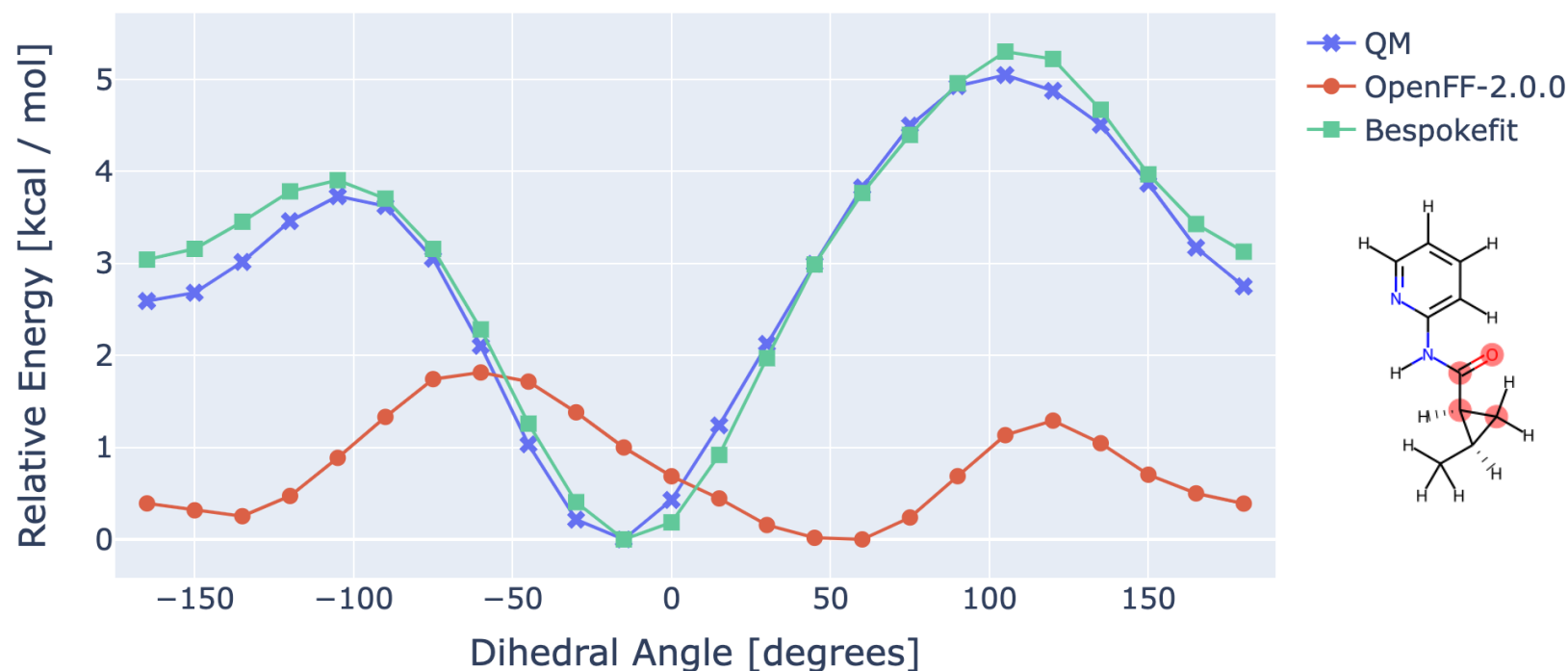
$$+ \sum_{\text{all torsions}} \sum_n K_\phi [1 - \cos(n\phi + \phi_0)]$$



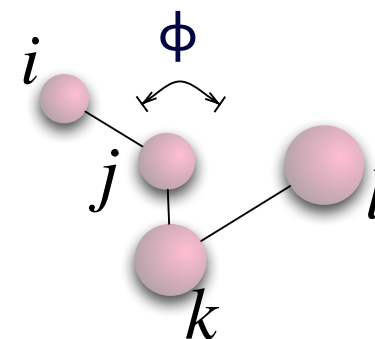
```
[ dihedrals ]
;ai    aj    ak    al    funct    phi    k
51     54     57     60     1      0.000000000000    1.796983533954    3
51     54     57     60     1      180.000000000000    1.064375944885    2
51     54     57     60     1      180.000000000000    3.655209445001    1
```

- Even more annoyingly, there are at least 3 different equivalent representations of dihedrals

# It can be difficult to get these right!

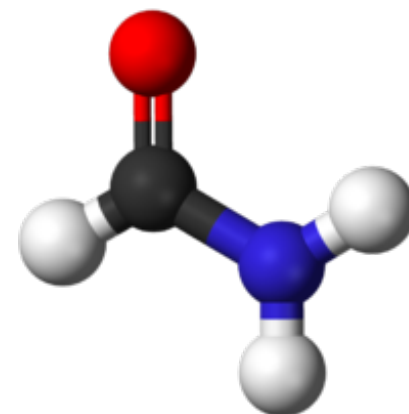
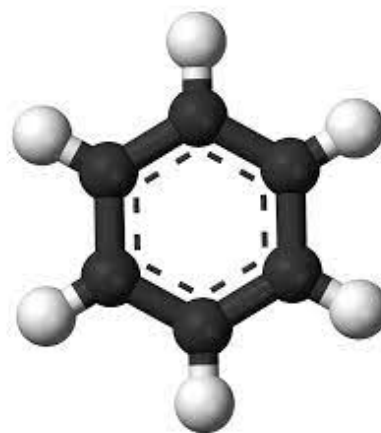
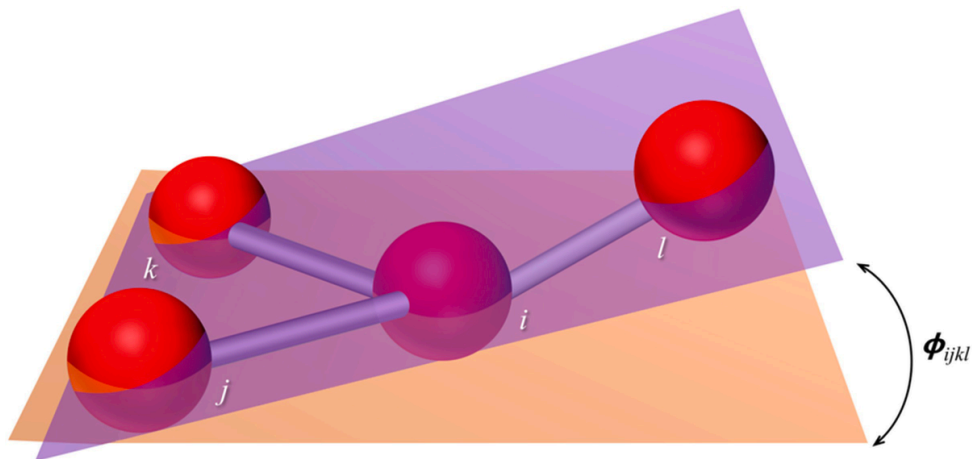


- NOTE: the energy of changing the torsion is not JUST the torsional potential, but involves the nonbonded interaction between atoms  $i$  and  $l$ .



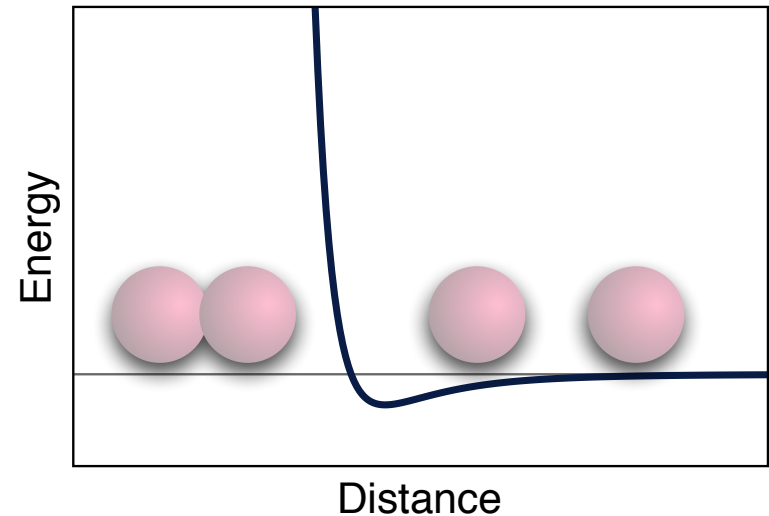
# Improper torsions

- A way to keep molecular systems with conjugated double bonds planar
- Energy minimum is at 0 degrees



# Lennard-Jones terms

$$U(r_{ij}) = 4\epsilon \left( \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right)$$



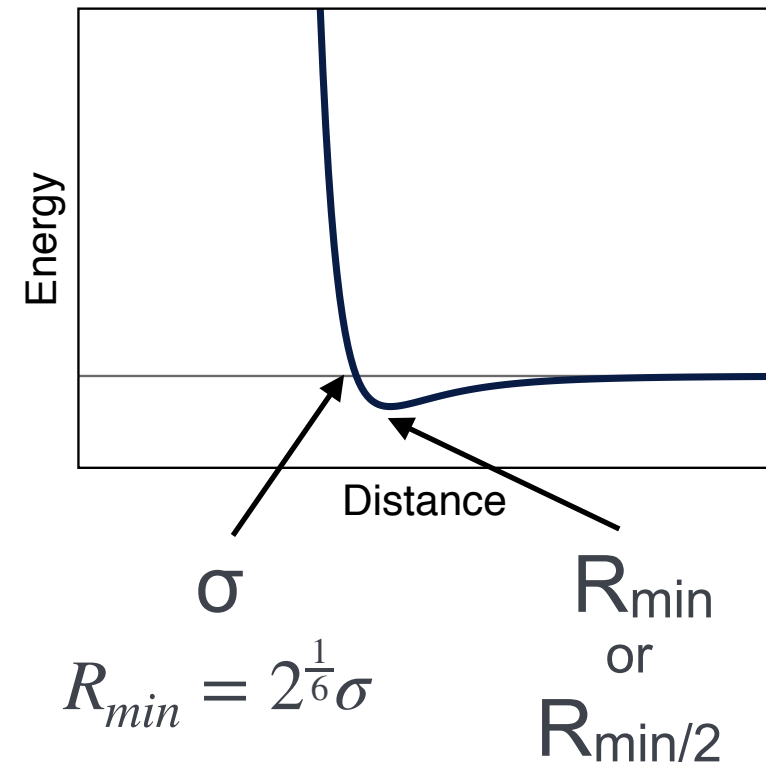
```
[ atomtypes ]  
; atom type name type      mass      charge      A      sigma      epsilon  
opls_111      OW      15.99940      0.000      A      3.12171e-01      7.94960e-01  
opls_112      HW      1.00800      0.000      A      0.00000e-01      0.00000e-01
```

- This just has sigma and epsilon for O and H. What about the O-H interaction???

# Lots of different, equivalent ways to write Lennard-Jones parameters!

$$U(r_{ij}) = 4\epsilon \left( \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right)$$

$$U(r_{ij}) = \frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6}$$



# Combining rules

```
[ defaults ]  
; nbfunc      comb-rule      gen-pairs      fudgeLJ  fudgeQQ  
1             3              yes              0.5      0.5
```

Geometric  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$   $\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$

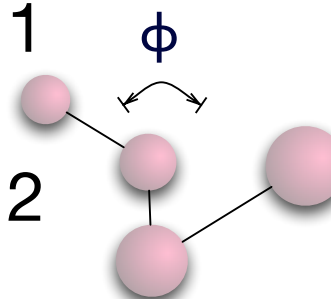
Lorentz-Berthelot  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$   $\sigma_{ij} = \frac{1}{2} (\sigma_i + \sigma_j)$

Works better than it should (but not perfect)



# Exclusions?

- What about atoms separated by no bonds (1,2)?
- What about atoms separated by 1 bond (1,3)?
- What about atoms separated by 2 bonds (1,4)?



```
[ defaults ]  
; nbfunc      comb-rule      gen-pairs      fudgeLJ  fudgeQQ  
1             3              yes              0.5      0.5
```

# Other terms you could put in:

- Hydrogen bonding terms
- Polarizability
  - The ability for charge distributions to adjust to nuclear positions
  - Still classical
- Other many body terms
  - charge penetration
  - Force field that break bonds (ReaxFF)
  - Machine learning terms (ANI multibody neural net potentials, many others.
- BUT pairwise nonbonded potentials are usually faster than general multibody terms for MD

# How do you choose force field parameters

- Bonded parameters:
  - Match to QM calculations
- Nonbonded parameters
  - Charges:
    - Match to electrostatics potential (RESP\_
    - AM1 with bond charge corrections (AM1-BCC)
- van der Waals
  - Match thermodynamic properties of liquids
    - Densities, enthalpies of vaporization or mixing + others

# What sorts of force fields are out there?

- OPLS-AA / OPLS-UA
  - Small molecules, proteins
  - Evolved out of Bill Jorgensen's work
  - Now run by Schrödinger
  - Biomolecules and small molecule ligands
- TRAPPE
  - Coming mostly from Ilja Siepmann's group
  - Designed for phase equilibrium

# Other Force Fields

- **CHARMM**

- Evolved from the work in Martin Karplus's group at Harvard over ~50 years.
- Alex MacKerrell (Maryland) and Benoit Roux (Chicago) are longest contributing developers but MANY more.
- Proteins, lipids, nucleic acids small molecules

- **AMBER**

- Evolved from the work in Peter Kollman's group at UCSF, and his academic descendants.
- GAFF small molecule force field (Junmei Wang)
- Proteins and nucleic acids

# Other Force Fields



- Open Force Field

- <http://www.openforcefield.org>
- New effort, focused on open software as well as open force fields
- Small molecule based, in the process of adding biopolymers
- Can build from SMILES, sdf, mol2
- Difference is: it matches molecular parameters in a more general way and DOESN'T use atom types

# Other Force Fields

- What about polymers?
  - OPLS often used.
  - COMPASS
  - MMFF
- What about metals?
  - Lots of metal specific functional forms.
  - Generally requires more complicated interactions than the simple nonbonded potentials

# Force Field Assignment Tools

- AmberTools/tleap
- CHARMM-GUI/CGENFF
- OpenFF toolkit
- foyer
- pysimm (polymers)
- OpenKIM (metals)



# But What Force Field Should I Use?

- The one that allows you to test the hypothesis that you want to

OR

- The one that is validated as an effective screening mechanism to predict the property of interest to other researchers.
- Has the force field I want to use been validated on sufficiently similar systems?
- "All models are wrong. Some models are useful."