
1 Force fields

1.1 Introduction

The term force field is slightly misleading, since it refers to the parameters of the potential used to calculate the forces (via gradient) in molecular dynamics simulations. The underlying idea is to create a certain number of atom types upon which any bonds, angles, impropers, dihedrals and long-range interactions may be described. More atomtypes than elements are necessary, since the chemical surroundings greatly influence the parameters. The angle between carbons in alkyne side chains will obviously differ from that of carbons in a benzene ring, so two different atom types are used to describe them. Since all simulations rely greatly upon the correctness of these parameters, it is important to know, where these parameters come from and how they may be derived in case some are lacking. One of the most commonly used force fields is the OPLS (Optimized Potential for Liquid Simulations) force field, developed by William L. Jorgensen at Purdue University and later at Yale University. It shall be our main example here.

1.2 A word on AMBER

The functional form of the OPLS force field is very similar to that of AMBER (Assisted Model Building and Energy Refinement), originally developed by Peter Kollman's group at the University of California San Francisco, which has the form [1, 2]:

$$\begin{aligned} V(\{\vec{r}_N\}) = & \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 \\ & + \sum_{torsions} \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)] \\ & + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} + \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \end{aligned}$$

1.3 Functional form of OPLS

Based upon AMBER, Jorgensen et al. created the OPLS force field using the following functional form [3, 4, 5, 6, 7]:

$$\begin{aligned} V(\{\vec{r}_N\}) = & \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 \\ & + \sum_{torsions} \left\{ \frac{V_1}{2} [1 + \cos(\phi)] + \frac{V_2}{2} [1 - \cos(2\phi)] + \frac{V_3}{2} [1 + \cos(3\phi)] + \frac{V_4}{2} [1 - \cos(4\phi)] \right\} \\ & + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} + \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j e^2}{r_{ij}} \right\} f_{ij} \end{aligned}$$

1.4 Derivation of the parameters

Bonds and angles are described by harmonic potentials, since they are very strong and fluctuate only slightly around their equilibrium values at room temperature. In many simulations, bonds and even angles are constrained to fixed values to save computational effort. The dihedral potential is described by a cosine expansion and may take any value within 360° depending on the height of the barrier between the low energy conformations, which makes the precision of the dihedral potential barrier crucial for many polymer properties. Dihedral potentials always possess a symmetry around 180° . The long range interactions are only counted for atoms three or more bonds apart. They consist of Coulomb and Lennard Jones two-body interaction terms. The Lennard Jones potential is a combination of attractive van der Waals forces due to dipole-dipole interactions and empirical repulsive forces due to Pauli repulsion. The scaling factor f_{ij} equals 0.5 for 1-4 interactions and one otherwise.

1.4 Derivation of the parameters

Since most molecular dynamics publications simply cite the force field used or give a list of parameters without further explanation, this section shall shed some light on how these parameters are found [1]. It should be noted, that in general any force field derivation must start with very small molecules, such as CH_3 , CH_4 , or maybe benzene, to limit the number of variables. Larger molecules can only be considered once the force field parameters for most constituents are considered to be known.

1.4.1 Bonds and angles

The equilibrium values for bonds as well as angles are usually simply taken from x-ray data. The values for the force constants are derived by fitting to experimental vibrational frequency data. While one should actually go through a more complex procedure to ensure that the geometries of simple molecules match experimental data as good as possible *after* energy minimization, it is believed that in most cases the difference is negligible.

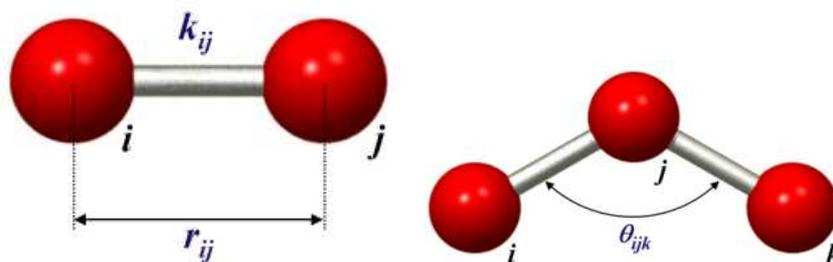


Fig. 1: Bond and angle between atoms i, j, k including corresponding force constants.

1.4 Derivation of the parameters

1.4.2 Dihedrals

First off, there are two different approaches as to how to calculate dihedral potentials. One is to optimize the dihedral potential for the simplest possible molecule and then apply it to larger ones containing the same dihedral, the other is to optimize the dihedral parameters to best describe a large number of different molecules. Note that, while the latter method might sound more accurate at first, it also leads to a dependence on the set of chosen molecules. In both cases, the dihedral parameters are computed from ab initio methods as follows [8]:

i. Ab initio calculations

- Scan dihedral (or improper) of interest
- Optimize geometry at each step
- Calculate change in potential energy

For this purpose, either perturbation (MP2), restricted Hartree-Fock (RHF) or hybrid methods between Hartree-Fock and density functional theory (B3LYP) are used. The basis sets chosen for the geometry optimization are at least 6-31g and may go up to 6-311g**, depending on the size of the molecule in question.

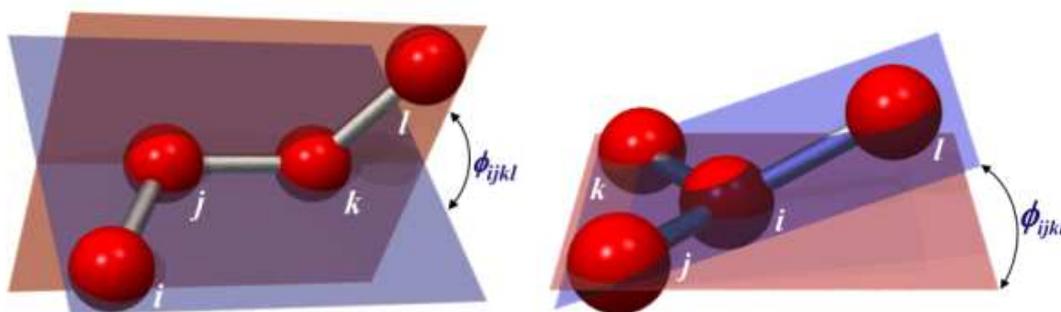


Fig. 2: Dihedral and improper between atoms i,j,k,l including corresponding force constants.

ii. Potential energy according to MD

- Set dihedral parameters to zero
- Compute potential energy of each optimized configuration

Note that this requires the knowledge of all other force field parameters in the molecule.

iii. Fitting of the parameters

- Subtract MD from ab initio results

1.4 Derivation of the parameters

- Obtain parameters from fitting a dihedral function to the resulting curve

The difference between ab initio and MD results corresponds exactly to the influence of the dihedral, since all other interactions are already included in the force field. The most common mathematical representations of dihedrals are:

- Proper dihedrals

$$V(\{\phi\}) = C_n \cdot (1 + \cos(n\phi))$$

- Ryckaert-Belleman dihedrals

$$V(\{\phi\}) = \sum_{n=0}^5 C_n \cos^n(\psi)$$

where $\phi = \psi - 180^\circ$.

iv. Checking the results

- Rerun MD simulations using new parameters
- Compare MD and ab initio results

1.4.3 Charges

To obtain the partial charges of a molecule in question, whose values will no longer be altered during the MD runs (unless a polarizable force field is used), it is important to make sure the charge calculations are done in the equilibrium, conformation. This may be achieved by optimizing the geometry using any ab initio method with a basis set better than or equal to 6-31G. The charges may then be obtained by fitting to the electrostatic potential, i.e. adjusting the partial charges at the centers of the nuclei in such a fashion that the electrostatic potential given by the wave functions is best reproduced. One of the favored methods is CHELPG (CHarges from ELectrostatic Potentials using a Grid based method) [9]. Charge calculations are often done using higher-level methods and basis sets than the geometry optimization, e.g. B3LYP/cc-pVTZ or HF/6-31G**. There are two known problems with this straight-forward approach however: Firstly, oftentimes considerable variation is seen when charges are computed for different conformations of a molecule, which is especially problematic for molecules with multiple low energy conformations such as propylamine [10]. Secondly, standard electrostatic potential (ESP) charges tend to miscalculate charges of "buried" atoms, since they are statistically underdetermined and often assume too large values for nonpolar atoms. For these reasons, the restricted ESP charge model was developed [11]. It consists of a least-squares fit of the charges to the electrostatic potential (as before), but with hyperbolic restraints on heavy atom charges. This is followed by a second fitting stage, needed to fit methyl groups which require equivalent charges on hydrogen atoms which are not equivalent by molecular symmetry.

1.5 Testing the force field

1.4.4 Van der Waals parameters

This is the area where the difference between OPLS/AMBER and other force fields comes into play: Jorgensen et al. pioneered on creating force field parameters for organic molecules focusing on systems explicitly taking the solvent into account. As a matter of fact, the OPLS/AMBER force field for peptides and proteins takes most bond, angle and dihedral parameters from the force field developed by Weiner et al [12]. This is achieved by carrying out Monte Carlo (MC) simulations of organic liquids, e.g. CH_4 or C_2H_6 , and then empirically adjusting the Lennard Jones (σ and ϵ) parameters to match the experimental densities and enthalpies of vaporization. A difficult issue is the factor to scale down the Lennard Jones 1-4 interactions. It has to be done, since otherwise the r^{-12} term would lead to unphysically high repulsions. The value chosen is somewhat arbitrary and force field dependent, however. The same holds true for the combination rules for the Lennard Jones parameters, i.e. the choice of LJ parameters for the interaction between two different atom types. There are two main competing methods:

- Lorentz-Bertelot

$$\begin{aligned}\sigma_{ij} &= \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \\ \epsilon_{ij} &= (\epsilon_{ii}\epsilon_{jj})^{1/2}\end{aligned}$$

- Geometrical average

$$\begin{aligned}\sigma_{ij} &= (\sigma_{ii}\sigma_{jj})^{1/2} \\ \epsilon_{ij} &= (\epsilon_{ii}\epsilon_{jj})^{1/2}\end{aligned}$$

The method should be chosen in accordance with the force field to be used.

1.5 Testing the force field

The validity of the force field can be tested in multiple ways. For proteins, it is important that the Ramachandran plots for the dihedral angles correspond to those obtained experimentally to ensure sampling of the correct secondary structures in simulations. For liquids, the radial distribution function as well as the density should be correctly reproduced. The radius of gyration and the persistence length are checked for polymer chains.

1.6 Some details for use with *Gromacs*

When one has chosen a force field to use with *Gromacs* it is important that some input parameters for the simulations are in accordance with this force field. All important settings may be found in the headers of the **itp** and **rtp** files. They are as follows:

1.6 Some details for use with *Gromacs*

1.6.1 itp-file

- **nbfunc**

type of non-bonded function to be used

- ★ 1: Lennard Jones (OPLS, AMBER)
- ★ 2: Buckingham

- **comb-rule**

combination rule to be used for non-bonded interactions

- ★ 1: Given C^6 and $C^N \Rightarrow C_{ij}^M = \sqrt{C_i^M C_j^M}$
- ★ 2: $\sigma_{ij} = \frac{1}{2}(\sigma_1 + \sigma_j)$ and $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ (AMBER)
- ★ 3: $\sigma_{ij} = \sqrt{\sigma_1 \sigma_j}$ and $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ (OPLS)

- **genpairs**

generate LJ 1-4 pairs?

- ★ yes (OPLS, AMBER)
- ★ no

Note: do not add [**pairtypes**] section, if genpairs = yes

- **fudgeLJ**

scaling factor for LJ 1-4 interactions

- ★ 0.5 (OPLS, AMBER)

- **fudgeQQ**

scaling factor for Coulomb 1-4 interactions

- ★ 0.8333 (AMBER)
- ★ 0.5 (OPLS-UA)
- ★ 1/8 (OPLS-AA)

1.6.2 rtp-file

- **bonds**

function used to describe bonded potential

- ★ 1: bond (OPLS, AMBER)
- ★ 2: G96 bond
- ★ 3: morse
- ★ 4: cubic

1.6 Some details for use with *Gromacs*

- ★ 5: connection
- ★ 6: harmonic potential
- ★ 7: FENE
- **angles**
function used to describe angular potential
 - ★ 1: angle (OPLS, AMBER)
 - ★ 2: G96 angle
 - ★ 3: quadratic angle
- **dihedrals**
function used to describe dihedral potential
 - ★ 1: proper
 - ★ 2: improper
 - ★ 3: Ryckaert-Belleman (OPLS, AMBER)
- **impropers**
function used to describe improper potential
 - ★ 1: proper (OPLS, AMBER)
 - ★ 2: improper
 - ★ 3: Ryckaert-Belleman
- **all_dihedrals**
generate all dihedrals?
 - ★ 0: only for heavy atoms
 - ★ 1: yes (OPLS, AMBER)
- **nrexcl**
number of excluded neighbors for non-bonded interactions
 - ★ 3 (default, should *not* be changed)
- **HH14**
generate 1-4 interaction pairs between hydrogens?
 - ★ 0: no
 - ★ 1: yes (OPLS, AMBER)
- **RemoveDih**
remove propers over the same bond as a improper?

REFERENCES

★ 0: no (OPLS, AMBER)

★ 1: yes

All above informations are based on *Gromacs* version 3.3.1.

References

- [1] W. Cornell, P. Cieplak, P.A. Kollmann et al., *J. Am. Chem. Soc.* **117**, 5179-5197 (1995).
- [2] Y. Duan, C. Wu, P. Kollmann et al., *J. Comp. Chem.* **24**, 1999-2012 (2003).
- [3] W.L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1657-1666 (1988).
- [4] W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225-11236 (1996).
- [5] W.L. Jorgensen, *Encyclopedia of Comp. Chem.*, 1986-1988 (1998).
- [6] W.L. Jorgensen, *Encyclopedia of Comp. Chem.*, 1754-1762 (1998).
- [7] W.L. Jorgensen and J. Tirado-Rives, *PNAS* **102**, 6665-6670 (2005).
- [8] A.D. Mackerell, *J. Comp. Chem.* **25**, 1584-1604 (2004).
- [9] C.M. Breneman, and K.B. Wiberg, *J. Comp. Chem.* **11**, 361-373 (1989).
- [10] W. Cornell, P. Cieplak, C. Bayly, and P.A. Kollmann, *J. Am. Chem. Soc.* **115**, 9620-9631 (1993).
- [11] C. Bayly, P. Cieplak, W. Cornell, and P.A. Kollmann, *J. Phys. Chem.* **97**, 10269-10280 (1993).
- [12] S.J. Weiner, P.A. Kollmann, P.J. Weiner et al., *J. Am. Chem. Soc.* **106**, 765-784 (1984).