GROMACS

Groningen Machine for Chemical Simulations

Reference Manual

Version 2016

Chapter 4

Interaction function and force fields

To accommodate the potential functions used in some popular force fields (see 4.10), GROMACS offers a choice of functions, both for non-bonded interaction and for dihedral interactions. They are described in the appropriate subsections.

The potential functions can be subdivided into three parts

1. *Non-bonded*: Lennard-Jones or Buckingham, and Coulomb or modified Coulomb. The non-bonded interactions are computed on the basis of a neighbor list (a list of non-bonded atoms within a certain radius), in which exclusions are already removed.
2. *Bonded*: covalent bond-stretching, angle-bending, improper dihedrals, and proper dihedrals. These are computed on the basis of fixed lists.
3. *Restraints*: position restraints, angle restraints, distance restraints, orientation restraints and dihedral restraints, all based on fixed lists.

4.1 Non-bonded interactions

Non-bonded interactions in GROMACS are pair-additive and centro-symmetric:

$$V(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N) = \sum_{i<j} V_{ij}(\boldsymbol{r}_{ij}); \tag{4.1}$$

$$\boldsymbol{F}_i = -\sum_j \frac{dV_{ij}(r_{ij})}{dr_{ij}} \frac{\boldsymbol{r}_{ij}}{r_{ij}} = -\boldsymbol{F}_j \tag{4.2}$$

The non-bonded interactions contain a repulsion term, a dispersion term, and a Coulomb term. The repulsion and dispersion term are combined in either the Lennard-Jones (or 6-12 interaction), or the Buckingham (or exp-6 potential). In addition, (partially) charged atoms act through the Coulomb term.

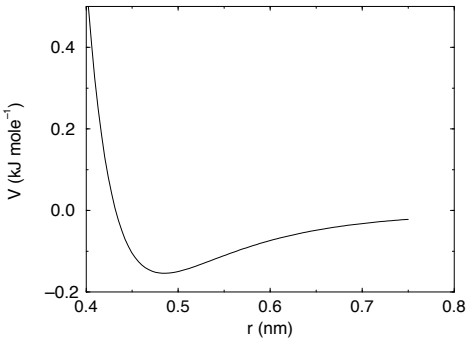

Figure 4.1: The Lennard-Jones interaction.

4.1.1 The Lennard-Jones interaction

The Lennard-Jones potential V_{LJ} between two atoms equals:

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \tag{4.3}$$

See also Fig. 4.1 The parameters $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ depend on pairs of *atom types*; consequently they are taken from a matrix of LJ-parameters. In the Verlet cut-off scheme, the potential is shifted by a constant such that it is zero at the cut-off distance.

The force derived from this potential is:

$$\boldsymbol{F}_i(\boldsymbol{r}_{ij}) = \left(12\, \frac{C_{ij}^{(12)}}{r_{ij}^{13}} - 6\, \frac{C_{ij}^{(6)}}{r_{ij}^{7}} \right) \frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.4}$$

The LJ potential may also be written in the following form:

$$V_{LJ}(\boldsymbol{r}_{ij}) = 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right) \tag{4.5}$$

In constructing the parameter matrix for the non-bonded LJ-parameters, two types of combination rules can be used within GROMACS, only geometric averages (type 1 in the input section of the force-field file):

$$\begin{aligned} C_{ij}^{(6)} &= \left(C_{ii}^{(6)}\, C_{jj}^{(6)} \right)^{1/2} \\ C_{ij}^{(12)} &= \left(C_{ii}^{(12)}\, C_{jj}^{(12)} \right)^{1/2} \end{aligned} \tag{4.6}$$

or, alternatively the Lorentz-Berthelot rules can be used. An arithmetic average is used to calculate σ_{ij}, while a geometric average is used to calculate ϵ_{ij} (type 2):

$$\begin{aligned} \sigma_{ij} &= \tfrac{1}{2}(\sigma_{ii} + \sigma_{jj}) \\ \epsilon_{ij} &= (\epsilon_{ii}\, \epsilon_{jj})^{1/2} \end{aligned} \tag{4.7}$$

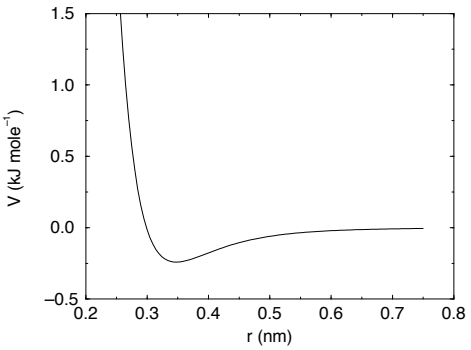

Figure 4.2: The Buckingham interaction.

finally an geometric average for both parameters can be used (type 3):

$$\begin{aligned} \sigma_{ij} &= (\sigma_{ii}\,\sigma_{jj})^{1/2} \\ \epsilon_{ij} &= (\epsilon_{ii}\,\epsilon_{jj})^{1/2} \end{aligned} \tag{4.8}$$

This last rule is used by the OPLS force field.

4.1.2 Buckingham potential

The Buckingham potential has a more flexible and realistic repulsion term than the Lennard-Jones interaction, but is also more expensive to compute. The potential form is:

$$V_{bh}(r_{ij}) = A_{ij}\exp(-B_{ij}r_{ij}) - \frac{C_{ij}}{r_{ij}^6} \tag{4.9}$$

See also Fig. 4.2. The force derived from this is:

$$\boldsymbol{F}_i(r_{ij}) = \left[A_{ij}B_{ij}\exp(-B_{ij}r_{ij}) - 6\frac{C_{ij}}{r_{ij}^7}\right]\frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.10}$$

4.1.3 Coulomb interaction

The Coulomb interaction between two charge particles is given by:

$$V_c(r_{ij}) = f\frac{q_i q_j}{\varepsilon_r r_{ij}} \tag{4.11}$$

See also Fig. 4.3, where $f = \frac{1}{4\pi\varepsilon_0} = 138.935\,458$ (see chapter 2)
The force derived from this potential is:

$$\boldsymbol{F}_i(\boldsymbol{r}_{ij}) = f\frac{q_i q_j}{\varepsilon_r r_{ij}^2}\frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.12}$$

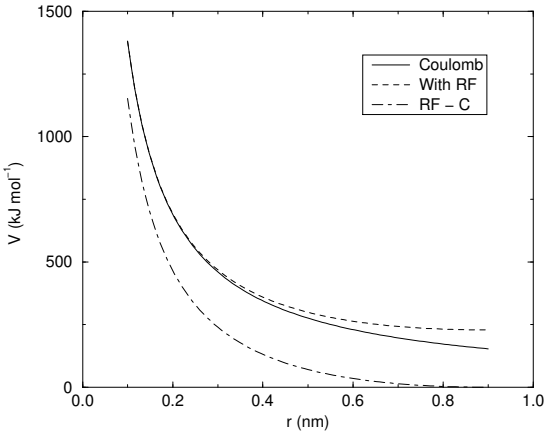

Figure 4.3: The Coulomb interaction (for particles with equal signed charge) with and without reaction field. In the latter case ε_r was 1, ε_{rf} was 78, and r_c was 0.9 nm. The dot-dashed line is the same as the dashed line, except for a constant.

A plain Coulomb interaction should only be used without cut-off or when all pairs fall within the cut-off, since there is an abrupt, large change in the force at the cut-off. In case you do want to use a cut-off, the potential can be shifted by a constant to make the potential the integral of the force. With the group cut-off scheme, this shift is only applied to non-excluded pairs. With the Verlet cut-off scheme, the shift is also applied to excluded pairs and self interactions, which makes the potential equivalent to a reaction field with $\varepsilon_{rf} = 1$ (see below).

In GROMACS the relative dielectric constant ε_r may be set in the in the input for `grompp`.

4.1.4 Coulomb interaction with reaction field

The Coulomb interaction can be modified for homogeneous systems by assuming a constant dielectric environment beyond the cut-off r_c with a dielectric constant of ε_{rf}. The interaction then reads:

$$V_{crf} = f\frac{q_i q_j}{\varepsilon_r r_{ij}}\left[1 + \frac{\varepsilon_{rf}-\varepsilon_r}{2\varepsilon_{rf}+\varepsilon_r}\frac{r_{ij}^3}{r_c^3}\right] - f\frac{q_i q_j}{\varepsilon_r r_c}\frac{3\varepsilon_{rf}}{2\varepsilon_{rf}+\varepsilon_r} \tag{4.13}$$

in which the constant expression on the right makes the potential zero at the cut-off r_c. For charged cut-off spheres this corresponds to neutralization with a homogeneous background charge. We can rewrite eqn. 4.13 for simplicity as

$$V_{crf} = f\frac{q_i q_j}{\varepsilon_r}\left[\frac{1}{r_{ij}} + k_{rf}\, r_{ij}^2 - c_{rf}\right] \tag{4.14}$$

with

$$k_{rf} = \frac{1}{r_c^3}\frac{\varepsilon_{rf}-\varepsilon_r}{(2\varepsilon_{rf}+\varepsilon_r)} \tag{4.15}$$

$$c_{rf} = \frac{1}{r_c} + k_{rf}\, r_c^2 = \frac{1}{r_c}\frac{3\varepsilon_{rf}}{(2\varepsilon_{rf}+\varepsilon_r)} \tag{4.16}$$

For large ε_{rf} the k_{rf} goes to $r_c^{-3}/2$, while for $\varepsilon_{rf} = \varepsilon_r$ the correction vanishes. In Fig. 4.3 the modified interaction is plotted, and it is clear that the derivative with respect to r_{ij} (= -force) goes to zero at the cut-off distance. The force derived from this potential reads:

$$\boldsymbol{F}_i(\boldsymbol{r}_{ij}) = f\frac{q_i q_j}{\varepsilon_r}\left[\frac{1}{r_{ij}^2} - 2k_{rf}r_{ij}\right]\frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.17}$$

The reaction-field correction should also be applied to all excluded atoms pairs, including self pairs, in which case the normal Coulomb term in eqns. 4.13 and 4.17 is absent.

Tironi *et al.* have introduced a generalized reaction field in which the dielectric continuum beyond the cut-off r_c also has an ionic strength I [75]. In this case we can rewrite the constants k_{rf} and c_{rf} using the inverse Debye screening length κ:

$$\kappa^2 = \frac{2I\,F^2}{\varepsilon_0\varepsilon_{rf}RT} = \frac{F^2}{\varepsilon_0\varepsilon_{rf}RT}\sum_{i=1}^{K} c_i z_i^2 \tag{4.18}$$

$$k_{rf} = \frac{1}{r_c^3}\frac{(\varepsilon_{rf}-\varepsilon_r)(1+\kappa r_c) + \frac{1}{2}\varepsilon_{rf}(\kappa r_c)^2}{(2\varepsilon_{rf}+\varepsilon_r)(1+\kappa r_c) + \varepsilon_{rf}(\kappa r_c)^2} \tag{4.19}$$

$$c_{rf} = \frac{1}{r_c}\frac{3\varepsilon_{rf}(1+\kappa r_c + \frac{1}{2}(\kappa r_c)^2)}{(2\varepsilon_{rf}+\varepsilon_r)(1+\kappa r_c) + \varepsilon_{rf}(\kappa r_c)^2} \tag{4.20}$$

where F is Faraday's constant, R is the ideal gas constant, T the absolute temperature, c_i the molar concentration for species i and z_i the charge number of species i where we have K different species. In the limit of zero ionic strength ($\kappa = 0$) eqns. 4.19 and 4.20 reduce to the simple forms of eqns. 4.15 and 4.16 respectively.

4.1.5 Modified non-bonded interactions

In GROMACS, the non-bonded potentials can be modified by a shift function, also called a force-switch function, since it switches the force to zero at the cut-off. The purpose of this is to replace the truncated forces by forces that are continuous and have continuous derivatives at the cut-off radius. With such forces the time integration produces smaller errors. But note that for Lennard-Jones interactions these errors are usually smaller than other errors, such as integration errors at the repulsive part of the potential. For Coulomb interactions we advise against using a shifted potential and for use of a reaction field or a proper long-range method such as PME.

There is *no* fundamental difference between a switch function (which multiplies the potential with a function) and a shift function (which adds a function to the force or potential) [76]. The switch function is a special case of the shift function, which we apply to the *force function* $F(r)$, related to the electrostatic or van der Waals force acting on particle i by particle j as:

$$\boldsymbol{F}_i = c\,F(r_{ij})\frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.21}$$

For pure Coulomb or Lennard-Jones interactions $F(r) = F_\alpha(r) = \alpha\, r^{-(\alpha+1)}$. The switched force

$F_s(r)$ can generally be written as:

$$\begin{aligned} F_s(r) &= F_\alpha(r) & r < r_1 \\ F_s(r) &= F_\alpha(r) + S(r) & r_1 \le r < r_c \\ F_s(r) &= 0 & r_c \le r \end{aligned} \tag{4.22}$$

When $r_1 = 0$ this is a traditional shift function, otherwise it acts as a switch function. The corresponding shifted potential function then reads:

$$V_s(r) = \int_r^\infty F_s(x)\,dx \tag{4.23}$$

The GROMACS **force switch** function $S_F(r)$ should be smooth at the boundaries, therefore the following boundary conditions are imposed on the switch function:

$$\begin{aligned} S_F(r_1) &= 0 \\ S'_F(r_1) &= 0 \\ S_F(r_c) &= -F_\alpha(r_c) \\ S'_F(r_c) &= -F'_\alpha(r_c) \end{aligned} \tag{4.24}$$

A 3^{rd} degree polynomial of the form

$$S_F(r) = A(r - r_1)^2 + B(r - r_1)^3 \tag{4.25}$$

fulfills these requirements. The constants A and B are given by the boundary condition at r_c:

$$\begin{aligned} A &= -\alpha \frac{(\alpha+4)r_c - (\alpha+1)r_1}{r_c^{\alpha+2}\,(r_c - r_1)^2} \\ B &= \alpha \frac{(\alpha+3)r_c - (\alpha+1)r_1}{r_c^{\alpha+2}\,(r_c - r_1)^3} \end{aligned} \tag{4.26}$$

Thus the total force function is:

$$F_s(r) = \frac{\alpha}{r^{\alpha+1}} + A(r - r_1)^2 + B(r - r_1)^3 \tag{4.27}$$

and the potential function reads:

$$V_s(r) = \frac{1}{r^\alpha} - \frac{A}{3}(r - r_1)^3 - \frac{B}{4}(r - r_1)^4 - C \tag{4.28}$$

where

$$C = \frac{1}{r_c^\alpha} - \frac{A}{3}(r_c - r_1)^3 - \frac{B}{4}(r_c - r_1)^4 \tag{4.29}$$

The GROMACS **potential-switch** function $S_V(r)$ scales the potential between r_1 and r_c, and has similar boundary conditions, intended to produce smoothly-varying potential and forces:

$$\begin{aligned} S_V(r_1) &= 1 \\ S'_V(r_1) &= 0 \\ S''_V(r_1) &= 0 \\ S_V(r_c) &= 0 \\ S'_V(r_c) &= 0 \\ S''_V(r_c) &= 0 \end{aligned} \tag{4.30}$$

The fifth-degree polynomial that has these properties is

$$S_V(r; r_1, r_c) = \frac{1 - 10(r - r_1)^3(r_c - r_1)^2 + 15(r - r_1)^4(r_c - r_1) - 6(r - r_1)}{(r_c - r_1)^5} \tag{4.31}$$

This implementation is found in several other simulation packages,[77, 78, 79] but differs from that in CHARMM.[80] Switching the potential leads to artificially large forces in the switching region, therefore it is not recommended to switch Coulomb interactions using this function,[76] but switching Lennard-Jones interactions using this function produces acceptable results.

4.1.6 Modified short-range interactions with Ewald summation

When Ewald summation or particle-mesh Ewald is used to calculate the long-range interactions, the short-range Coulomb potential must also be modified. Here the potential is switched to (nearly) zero at the cut-off, instead of the force. In this case the short range potential is given by:

$$V(r) = f\frac{\text{erfc}(\beta r_{ij})}{r_{ij}} q_i q_j, \tag{4.32}$$

where β is a parameter that determines the relative weight between the direct space sum and the reciprocal space sum and $\text{erfc}(x)$ is the complementary error function. For further details on long-range electrostatics, see sec. 4.8.

4.2 Bonded interactions

Bonded interactions are based on a fixed list of atoms. They are not exclusively pair interactions, but include 3- and 4-body interactions as well. There are *bond stretching* (2-body), *bond angle* (3-body), and *dihedral angle* (4-body) interactions. A special type of dihedral interaction (called *improper dihedral*) is used to force atoms to remain in a plane or to prevent transition to a configuration of opposite chirality (a mirror image).

4.2.1 Bond stretching

Harmonic potential

The bond stretching between two covalently bonded atoms i and j is represented by a harmonic potential:

$$V_b\left(r_{ij}\right) = \frac{1}{2} k_{ij}^b (r_{ij} - b_{ij})^2 \tag{4.33}$$

See also Fig. 4.4, with the force given by:

$$\boldsymbol{F}_i(\boldsymbol{r}_{ij}) = k_{ij}^b (r_{ij} - b_{ij}) \frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.34}$$

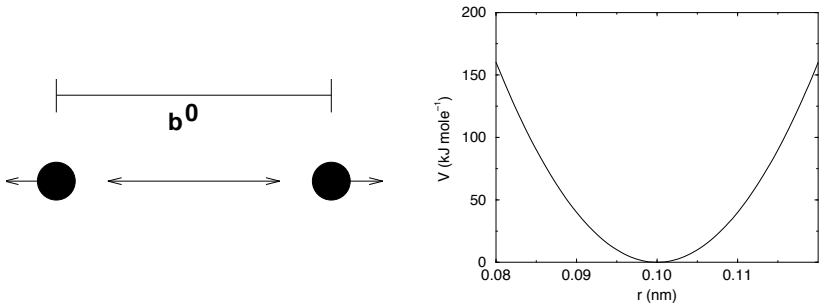

Figure 4.4: Principle of bond stretching (left), and the bond stretching potential (right).

Fourth power potential

In the GROMOS-96 force field [81], the covalent bond potential is, for reasons of computational efficiency, written as:

$$V_b\left(r_{ij}\right) = \frac{1}{4}k_{ij}^b\left(r_{ij}^2 - b_{ij}^2\right)^2 \tag{4.35}$$

The corresponding force is:

$$\boldsymbol{F}_i(\boldsymbol{r}_{ij}) = k_{ij}^b(r_{ij}^2 - b_{ij}^2)\,\boldsymbol{r}_{ij} \tag{4.36}$$

The force constants for this form of the potential are related to the usual harmonic force constant $k^{b,\mathrm{harm}}$ (sec. 4.2.1) as

$$2k^b b_{ij}^2 = k^{b,\mathrm{harm}} \tag{4.37}$$

The force constants are mostly derived from the harmonic ones used in GROMOS-87 [82]. Although this form is computationally more efficient (because no square root has to be evaluated), it is conceptually more complex. One particular disadvantage is that since the form is not harmonic, the average energy of a single bond is not equal to $\frac{1}{2}kT$ as it is for the normal harmonic potential.

4.2.2 Morse potential bond stretching

For some systems that require an anharmonic bond stretching potential, the Morse potential [83] between two atoms *i* and *j* is available in GROMACS. This potential differs from the harmonic potential in that it has an asymmetric potential well and a zero force at infinite distance. The functional form is:

$$V_{morse}(r_{ij}) = D_{ij}[1 - \exp(-\beta_{ij}(r_{ij} - b_{ij}))]^2, \tag{4.38}$$

See also Fig. 4.5, and the corresponding force is:

$$\begin{aligned}\mathbf{F}_{morse}(\mathbf{r}_{ij}) \quad = \quad & 2D_{ij}\beta_{ij}\exp(-\beta_{ij}(r_{ij} - b_{ij}))* \\ & [1 - \exp(-\beta_{ij}(r_{ij} - b_{ij}))]\frac{\mathbf{r}_{ij}}{r_{ij}},\end{aligned} \tag{4.39}$$

where D_{ij} is the depth of the well in kJ/mol, β_{ij} defines the steepness of the well (in nm^{-1}), and b_{ij} is the equilibrium distance in nm. The steepness parameter β_{ij} can be expressed in terms of

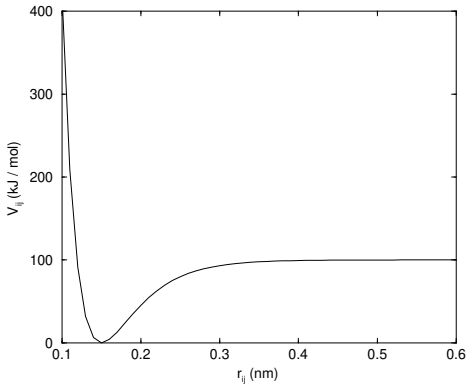

Figure 4.5: The Morse potential well, with bond length 0.15 nm.

the reduced mass of the atoms *i* and *j*, the fundamental vibration frequency ω_{ij} and the well depth D_{ij}:

$$\beta_{ij} = \omega_{ij}\sqrt{\frac{\mu_{ij}}{2D_{ij}}} \tag{4.40}$$

and because $\omega = \sqrt{k/\mu}$, one can rewrite β_{ij} in terms of the harmonic force constant k_{ij}:

$$\beta_{ij} = \sqrt{\frac{k_{ij}}{2D_{ij}}} \tag{4.41}$$

For small deviations $(r_{ij} - b_{ij})$, one can approximate the exp-term to first-order using a Taylor expansion:

$$\exp(-x) \approx 1 - x \tag{4.42}$$

and substituting eqn. 4.41 and eqn. 4.42 in the functional form:

$$\begin{aligned} V_{morse}(r_{ij}) &= D_{ij}[1 - \exp(-\beta_{ij}(r_{ij} - b_{ij}))]^2 \\ &= D_{ij}[1 - (1 - \sqrt{\tfrac{k_{ij}}{2D_{ij}}}(r_{ij} - b_{ij}))]^2 \\ &= \tfrac{1}{2}k_{ij}(r_{ij} - b_{ij}))^2 \end{aligned} \tag{4.43}$$

we recover the harmonic bond stretching potential.

4.2.3 Cubic bond stretching potential

Another anharmonic bond stretching potential that is slightly simpler than the Morse potential adds a cubic term in the distance to the simple harmonic form:

$$V_b\,(r_{ij}) = k^b_{ij}(r_{ij} - b_{ij})^2 + k^b_{ij}k^{cub}_{ij}(r_{ij} - b_{ij})^3 \tag{4.44}$$

A flexible water model (based on the SPC water model [84]) including a cubic bond stretching potential for the O-H bond was developed by Ferguson [85]. This model was found to yield a

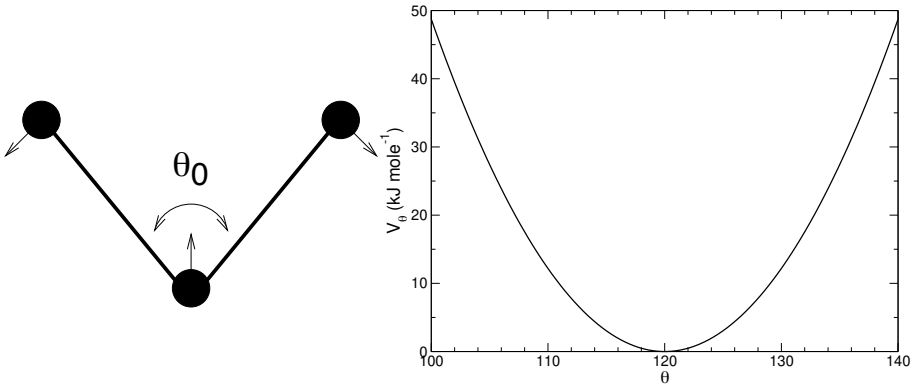

Figure 4.6: Principle of angle vibration (left) and the bond angle potential (right).

reasonable infrared spectrum. The Ferguson water model is available in the GROMACS library (`flexwat-ferguson.itp`). It should be noted that the potential is asymmetric: overstretching leads to infinitely low energies. The integration timestep is therefore limited to 1 fs.

The force corresponding to this potential is:

$$\boldsymbol{F}_i(\boldsymbol{r}_{ij}) = 2k^b_{ij}(r_{ij} - b_{ij})\,\frac{\boldsymbol{r}_{ij}}{r_{ij}} + 3k^b_{ij}k^{cub}_{ij}(r_{ij} - b_{ij})^2\,\frac{\boldsymbol{r}_{ij}}{r_{ij}} \tag{4.45}$$

4.2.4 FENE bond stretching potential

In coarse-grained polymer simulations the beads are often connected by a FENE (finitely extensible nonlinear elastic) potential [86]:

$$V_{\text{FENE}}(r_{ij}) = -\frac{1}{2}k^b_{ij}b^2_{ij}\log\left(1 - \frac{r^2_{ij}}{b^2_{ij}}\right) \tag{4.46}$$

The potential looks complicated, but the expression for the force is simpler:

$$F_{\text{FENE}}(\boldsymbol{r}_{ij}) = -k^b_{ij}\left(1 - \frac{r^2_{ij}}{b^2_{ij}}\right)^{-1}\boldsymbol{r}_{ij} \tag{4.47}$$

At short distances the potential asymptotically goes to a harmonic potential with force constant k^b, while it diverges at distance b.

4.2.5 Harmonic angle potential

The bond-angle vibration between a triplet of atoms i - j - k is also represented by a harmonic potential on the angle θ_{ijk}

$$V_a(\theta_{ijk}) = \frac{1}{2}k^\theta_{ijk}(\theta_{ijk} - \theta^0_{ijk})^2 \tag{4.48}$$

As the bond-angle vibration is represented by a harmonic potential, the form is the same as the bond stretching (Fig. 4.4).

The force equations are given by the chain rule:

$$\begin{aligned} \boldsymbol{F}_i &= -\frac{dV_a(\theta_{ijk})}{d\boldsymbol{r}_i} \\ \boldsymbol{F}_k &= -\frac{dV_a(\theta_{ijk})}{d\boldsymbol{r}_k} \\ \boldsymbol{F}_j &= -\boldsymbol{F}_i - \boldsymbol{F}_k \end{aligned} \quad \text{where} \quad \theta_{ijk} = \arccos \frac{(\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj})}{r_{ij} r_{kj}} \tag{4.49}$$

The numbering i, j, k is in sequence of covalently bonded atoms. Atom j is in the middle; atoms i and k are at the ends (see Fig. 4.6). **Note** that in the input in topology files, angles are given in degrees and force constants in kJ/mol/rad^2.

4.2.6 Cosine based angle potential

In the GROMOS-96 force field a simplified function is used to represent angle vibrations:

$$V_a(\theta_{ijk}) = \frac{1}{2} k^{\theta}_{ijk} \left(\cos(\theta_{ijk}) - \cos(\theta^0_{ijk}) \right)^2 \tag{4.50}$$

where

$$\cos(\theta_{ijk}) = \frac{\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj}}{r_{ij} r_{kj}} \tag{4.51}$$

The corresponding force can be derived by partial differentiation with respect to the atomic positions. The force constants in this function are related to the force constants in the harmonic form $k^{\theta,\mathrm{harm}}$ (4.2.5) by:

$$k^{\theta} \sin^2(\theta^0_{ijk}) = k^{\theta,\mathrm{harm}} \tag{4.52}$$

In the GROMOS-96 manual there is a much more complicated conversion formula which is temperature dependent. The formulas are equivalent at 0 K and the differences at 300 K are on the order of 0.1 to 0.2%. **Note** that in the input in topology files, angles are given in degrees and force constants in kJ/mol.

4.2.7 Restricted bending potential

The restricted bending (ReB) potential [87] prevents the bending angle θ from reaching the 180° value. In this way, the numerical instabilities due to the calculation of the torsion angle and potential are eliminated when performing coarse-grained molecular dynamics simulations.

To systematically hinder the bending angles from reaching the 180° value, the bending potential 4.50 is divided by a $\sin^2\theta$ factor:

$$V_{\mathrm{ReB}}(\theta_i) = \frac{1}{2} k_\theta \frac{(\cos\theta_i - \cos\theta_0)^2}{\sin^2\theta_i}. \tag{4.53}$$

Figure Fig. 4.7 shows the comparison between the ReB potential, 4.53, and the standard one 4.50. The wall of the ReB potential is very repulsive in the region close to 180° and, as a result, the bending angles are kept within a safe interval, far from instabilities. The power 2 of $\sin\theta_i$ in the denominator has been chosen to guarantee this behavior and allows an elegant differentiation:

$$F_{\mathrm{ReB}}(\theta_i) = \frac{2k_\theta}{\sin^4\theta_i} (\cos\theta_i - \cos\theta_0)(1 - \cos\theta_i \cos\theta_0) \frac{\partial \cos\theta_i}{\partial \vec{r}_k}. \tag{4.54}$$

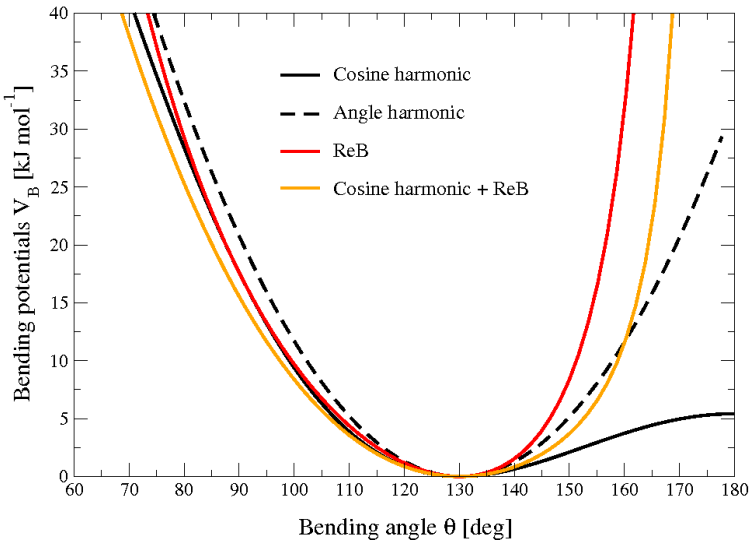

Figure 4.7: Bending angle potentials: cosine harmonic (solid black line), angle harmonic (dashed black line) and restricted bending (red) with the same bending constant $k_\theta = 85$ kJ mol^{-1} and equilibrium angle $\theta_0 = 130°$. The orange line represents the sum of a cosine harmonic ($k = 50$ kJ mol^{-1}) with a restricted bending ($k = 25$ kJ mol^{-1}) potential, both with $\theta_0 = 130°$.

Due to its construction, the restricted bending potential cannot be used for equilibrium θ_0 values too close to $0°$ or $180°$ (from experience, at least $10°$ difference is recommended). It is very important that, in the starting configuration, all the bending angles have to be in the safe interval to avoid initial instabilities. This bending potential can be used in combination with any form of torsion potential. It will always prevent three consecutive particles from becoming collinear and, as a result, any torsion potential will remain free of singularities. It can be also added to a standard bending potential to affect the angle around $180°$, but to keep its original form around the minimum (see the orange curve in Fig. 4.7).

4.2.8 Urey-Bradley potential

The Urey-Bradley bond-angle vibration between a triplet of atoms i - j - k is represented by a harmonic potential on the angle θ_{ijk} and a harmonic correction term on the distance between the atoms i and k. Although this can be easily written as a simple sum of two terms, it is convenient to have it as a single entry in the topology file and in the output as a separate energy term. It is used mainly in the CHARMm force field [88]. The energy is given by:

$$V_a(\theta_{ijk}) = \frac{1}{2}k^{\theta}_{ijk}(\theta_{ijk} - \theta^0_{ijk})^2 + \frac{1}{2}k^{UB}_{ijk}(r_{ik} - r^0_{ik})^2 \tag{4.55}$$

The force equations can be deduced from sections 4.2.1 and 4.2.5.

4.2.9 Bond-Bond cross term

The bond-bond cross term for three particles i, j, k forming bonds $i-j$ and $k-j$ is given by [89]:

$$V_{rr'} = k_{rr'} \left(|\boldsymbol{r}_i - \boldsymbol{r}_j| - r_{1e}\right) \left(|\boldsymbol{r}_k - \boldsymbol{r}_j| - r_{2e}\right) \tag{4.56}$$

where $k_{rr'}$ is the force constant, and r_{1e} and r_{2e} are the equilibrium bond lengths of the $i-j$ and $k-j$ bonds respectively. The force associated with this potential on particle i is:

$$\boldsymbol{F}_i = -k_{rr'} \left(|\boldsymbol{r}_k - \boldsymbol{r}_j| - r_{2e}\right) \frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \tag{4.57}$$

The force on atom k can be obtained by swapping i and k in the above equation. Finally, the force on atom j follows from the fact that the sum of internal forces should be zero: $\boldsymbol{F}_j = -\boldsymbol{F}_i - \boldsymbol{F}_k$.

4.2.10 Bond-Angle cross term

The bond-angle cross term for three particles i, j, k forming bonds $i-j$ and $k-j$ is given by [89]:

$$V_{r\theta} = k_{r\theta} \left(|\boldsymbol{r}_i - \boldsymbol{r}_k| - r_{3e}\right) \left(|\boldsymbol{r}_i - \boldsymbol{r}_j| - r_{1e} + |\boldsymbol{r}_k - \boldsymbol{r}_j| - r_{2e}\right) \tag{4.58}$$

where $k_{r\theta}$ is the force constant, r_{3e} is the $i-k$ distance, and the other constants are the same as in Equation 4.56. The force associated with the potential on atom i is:

$$\boldsymbol{F}_i = -k_{r\theta} \left[\left(|\boldsymbol{r}_i - \boldsymbol{r}_k| - r_{3e}\right) \frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \left(|\boldsymbol{r}_i - \boldsymbol{r}_j| - r_{1e} + |\boldsymbol{r}_k - \boldsymbol{r}_j| - r_{2e}\right) \frac{\boldsymbol{r}_i - \boldsymbol{r}_k}{|\boldsymbol{r}_i - \boldsymbol{r}_k|} \right] \tag{4.59}$$

4.2.11 Quartic angle potential

For special purposes there is an angle potential that uses a fourth order polynomial:

$$V_q(\theta_{ijk}) = \sum_{n=0}^{5} C_n (\theta_{ijk} - \theta_{ijk}^0)^n \tag{4.60}$$

4.2.12 Improper dihedrals

Improper dihedrals are meant to keep planar groups (*e.g.* aromatic rings) planar, or to prevent molecules from flipping over to their mirror images, see Fig. 4.8.

Improper dihedrals: harmonic type

The simplest improper dihedral potential is a harmonic potential; it is plotted in Fig. 4.9.

$$V_{id}(\xi_{ijkl}) = \frac{1}{2} k_\xi (\xi_{ijkl} - \xi_0)^2 \tag{4.61}$$

Since the potential is harmonic it is discontinuous, but since the discontinuity is chosen at 180° distance from ξ_0 this will never cause problems. **Note** that in the input in topology files, angles are given in degrees and force constants in kJ/mol/rad^2.

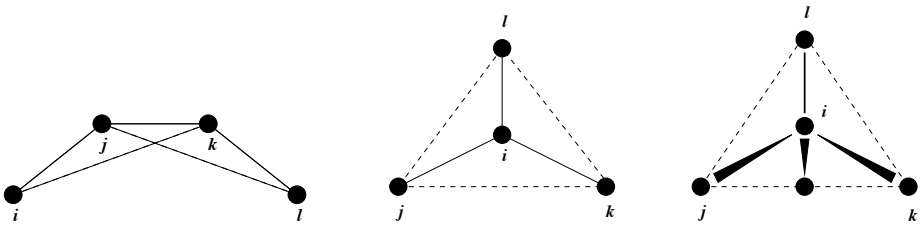

Figure 4.8: Principle of improper dihedral angles. Out of plane bending for rings (left), substituents of rings (middle), out of tetrahedral (right). The improper dihedral angle ξ is defined as the angle between planes (i,j,k) and (j,k,l) in all cases.

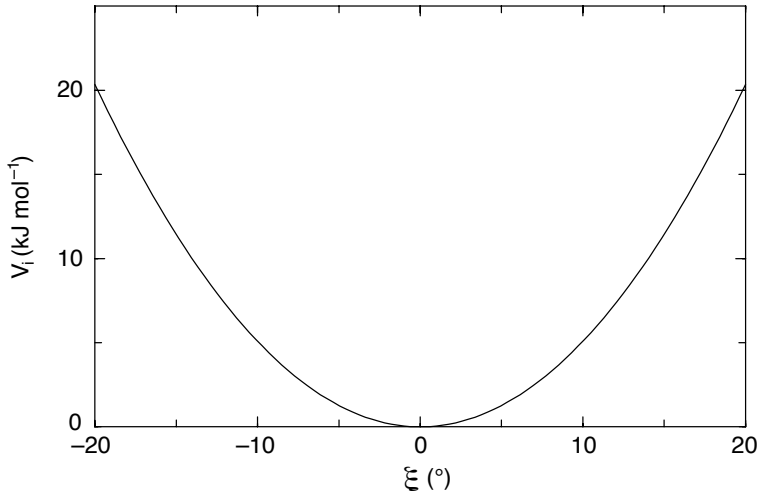

Figure 4.9: Improper dihedral potential.

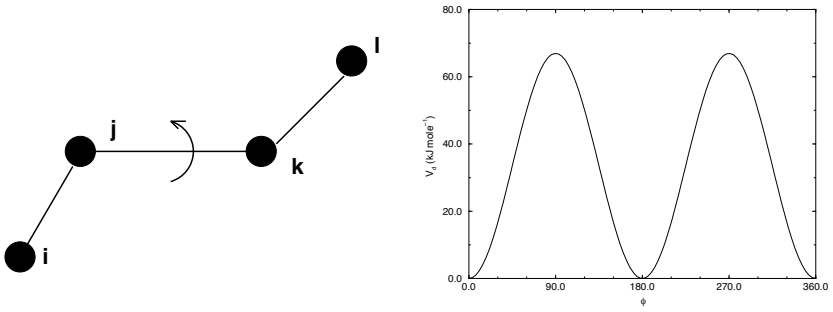

Figure 4.10: Principle of proper dihedral angle (left, in *trans* form) and the dihedral angle potential (right).

Improper dihedrals: periodic type

This potential is identical to the periodic proper dihedral (see below). There is a separate dihedral type for this (type 4) only to be able to distinguish improper from proper dihedrals in the parameter section and the output.

4.2.13 Proper dihedrals

For the normal dihedral interaction there is a choice of either the GROMOS periodic function or a function based on expansion in powers of $\cos\phi$ (the so-called Ryckaert-Bellemans potential). This choice has consequences for the inclusion of special interactions between the first and the fourth atom of the dihedral quadruple. With the periodic GROMOS potential a special 1-4 LJ-interaction must be included; with the Ryckaert-Bellemans potential *for alkanes* the 1-4 interactions must be excluded from the non-bonded list. **Note:** Ryckaert-Bellemans potentials are also used in *e.g.* the OPLS force field in combination with 1-4 interactions. You should therefore not modify topologies generated by `pdb2gmx` in this case.

Proper dihedrals: periodic type

Proper dihedral angles are defined according to the IUPAC/IUB convention, where ϕ is the angle between the ijk and the jkl planes, with **zero** corresponding to the *cis* configuration (i and l on the same side). There are two dihedral function types in GROMACS topology files. There is the standard type 1 which behaves like any other bonded interactions. For certain force fields, type 9 is useful. Type 9 allows multiple potential functions to be applied automatically to a single dihedral in the `[ dihedral ]` section when multiple parameters are defined for the same atomtypes in the `[ dihedraltypes ]` section.

$$V_d(\phi_{ijkl}) = k_\phi(1 + \cos(n\phi - \phi_s)) \tag{4.62}$$

C_0	9.28	C_2	-13.12	C_4	26.24
C_1	12.16	C_3	-3.06	C_5	-31.5

Table 4.1: Constants for Ryckaert-Bellemans potential (kJ mol^{-1}).

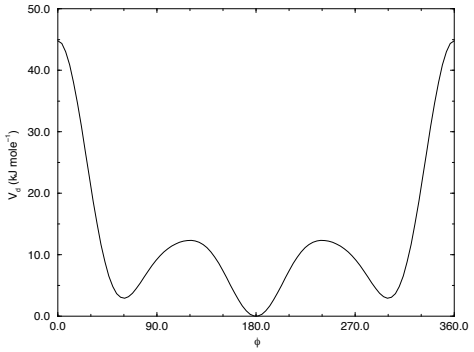

Figure 4.11: Ryckaert-Bellemans dihedral potential.

Proper dihedrals: Ryckaert-Bellemans function

For alkanes, the following proper dihedral potential is often used (see Fig. 4.11):

$$V_{rb}(\phi_{ijkl}) = \sum_{n=0}^{5} C_n(\cos(\psi))^n, \tag{4.63}$$

where $\psi = \phi - 180°$.
Note: A conversion from one convention to another can be achieved by multiplying every coefficient C_n by $(-1)^n$.

An example of constants for C is given in Table 4.1.

(**Note:** The use of this potential implies exclusion of LJ interactions between the first and the last atom of the dihedral, and ψ is defined according to the "polymer convention" ($\psi_{trans} = 0$).)

The RB dihedral function can also be used to include Fourier dihedrals (see below):

$$V_{rb}(\phi_{ijkl}) = \frac{1}{2}\left[F_1(1+\cos(\phi)) + F_2(1-\cos(2\phi)) + F_3(1+\cos(3\phi)) + F_4(1-\cos(4\phi))\right] \tag{4.64}$$

Because of the equalities $\cos(2\phi) = 2\cos^2(\phi) - 1$, $\cos(3\phi) = 4\cos^3(\phi) - 3\cos(\phi)$ and $\cos(4\phi) = 8\cos^4(\phi) - 8\cos^2(\phi) + 1$ one can translate the OPLS parameters to Ryckaert-Bellemans param-

eters as follows:

$$\begin{aligned} C_0 &= F_2 + \tfrac{1}{2}(F_1 + F_3) \\ C_1 &= \tfrac{1}{2}(-F_1 + 3\,F_3) \\ C_2 &= -F_2 + 4\,F_4 \\ C_3 &= -2\,F_3 \\ C_4 &= -4\,F_4 \\ C_5 &= 0 \end{aligned} \tag{4.65}$$

with OPLS parameters in protein convention and RB parameters in polymer convention (this yields a minus sign for the odd powers of $\cos(\phi)$).
Note: Mind the conversion from **kcal mol**$^{-1}$ for literature OPLS and RB parameters to **kJ mol**$^{-1}$ in GROMACS.

Proper dihedrals: Fourier function

The OPLS potential function is given as the first three or four [90] cosine terms of a Fourier series. In GROMACS the four term function is implemented:

$$V_F(\phi_{ijkl}) = \frac{1}{2}\left[C_1(1+\cos(\phi)) + C_2(1-\cos(2\phi)) + C_3(1+\cos(3\phi)) + C_4(1-\cos(4\phi))\right], \tag{4.66}$$

Internally, GROMACS uses the Ryckaert-Bellemans code to compute Fourier dihedrals (see above), because this is more efficient.
Note: Mind the conversion from *k*cal mol^{-1} for literature OPLS parameters to **kJ mol**$^{-1}$ in GROMACS.

Proper dihedrals: Restricted torsion potential

In a manner very similar to the restricted bending potential (see 4.2.7), a restricted torsion/dihedral potential is introduced:

$$V_{\mathrm{ReT}}(\phi_i) = \frac{1}{2}k_\phi \frac{(\cos\phi_i - \cos\phi_0)^2}{\sin^2\phi_i} \tag{4.67}$$

with the advantages of being a function of $\cos\phi$ (no problems taking the derivative of $\sin\phi$) and of keeping the torsion angle at only one minimum value. In this case, the factor $\sin^2\phi$ does not allow the dihedral angle to move from the $[-180°{:}0]$ to $[0{:}180°]$ interval, i.e. it cannot have maxima both at $-\phi_0$ and $+\phi_0$ maxima, but only one of them. For this reason, all the dihedral angles of the starting configuration should have their values in the desired angles interval and the the equilibrium ϕ_0 value should not be too close to the interval limits (as for the restricted bending potential, described in 4.2.7, at least $10°$ difference is recommended).

Proper dihedrals: Combined bending-torsion potential

When the four particles forming the dihedral angle become collinear (this situation will never happen in atomistic simulations, but it can occur in coarse-grained simulations) the calculation of

the torsion angle and potential leads to numerical instabilities. One way to avoid this is to use the restricted bending potential (see 4.2.7) that prevents the dihedral from reaching the 180° value.

Another way is to disregard any effects of the dihedral becoming ill-defined, keeping the dihedral force and potential calculation continuous in entire angle range by coupling the torsion potential (in a cosine form) with the bending potentials of the adjacent bending angles in a unique expression:

$$V_{\mathrm{CBT}}(\theta_{i-1}, \theta_i, \phi_i) = k_\phi \sin^3\theta_{i-1} \sin^3\theta_i \sum_{n=0}^{4} a_n \cos^n\phi_i. \tag{4.68}$$

This combined bending-torsion (CBT) potential has been proposed by [91] for polymer melt simulations and is extensively described in [87].

This potential has two main advantages:

- it does not only depend on the dihedral angle ϕ_i (between the $i-2$, $i-1$, i and $i+1$ beads) but also on the bending angles θ_{i-1} and θ_i defined from three adjacent beads ($i-2$, $i-1$ and i, and $i-1$, i and $i+1$, respectively). The two $\sin^3\theta$ pre-factors, tentatively suggested by [92] and theoretically discussed by [93], cancel the torsion potential and force when either of the two bending angles approaches the value of 180°.
- its dependence on ϕ_i is expressed through a polynomial in $\cos\phi_i$ that avoids the singularities in $\phi = 0^\circ$ or 180° in calculating the torsional force.

These two properties make the CBT potential well-behaved for MD simulations with weak constraints on the bending angles or even for steered / non-equilibrium MD in which the bending and torsion angles suffer major modifications. When using the CBT potential, the bending potentials for the adjacent θ_{i-1} and θ_i may have any form. It is also possible to leave out the two angle bending terms (θ_{i-1} and θ_i) completely. Fig. 4.12 illustrates the difference between a torsion potential with and without the $\sin^3\theta$ factors (blue and gray curves, respectively). Additionally, the derivative of V_{CBT} with respect to the Cartesian variables is straightforward:

$$\frac{\partial V_{\mathrm{CBT}}(\theta_{i-1}, \theta_i, \phi_i)}{\partial \vec{r}_l} = \frac{\partial V_{\mathrm{CBT}}}{\partial \theta_{i-1}} \frac{\partial \theta_{i-1}}{\partial \vec{r}_l} + \frac{\partial V_{\mathrm{CBT}}}{\partial \theta_i} \frac{\partial \theta_i}{\partial \vec{r}_l} + \frac{\partial V_{\mathrm{CBT}}}{\partial \phi_i} \frac{\partial \phi_i}{\partial \vec{r}_l} \tag{4.69}$$

The CBT is based on a cosine form without multiplicity, so it can only be symmetrical around 0°. To obtain an asymmetrical dihedral angle distribution (e.g. only one maximum in $[-180^\circ{:}180^\circ]$ interval), a standard torsion potential such as harmonic angle or periodic cosine potentials should be used instead of a CBT potential. However, these two forms have the inconveniences of the force derivation ($1/\sin\phi$) and of the alignment of beads (θ_i or $\theta_{i-1} = 0^\circ, 180^\circ$). Coupling such non-$\cos\phi$ potentials with $\sin^3\theta$ factors does not improve simulation stability since there are cases in which θ and ϕ are simultaneously 180°. The integration at this step would be possible (due to the cancelling of the torsion potential) but the next step would be singular (θ is not 180° and ϕ is very close to 180°).

4.2.14 Tabulated bonded interaction functions

For full flexibility, any functional shape can be used for bonds, angles and dihedrals through user-supplied tabulated functions. The functional shapes are:

$$V_b(r_{ij}) \quad = \quad k\, f_n^b(r_{ij}) \tag{4.70}$$

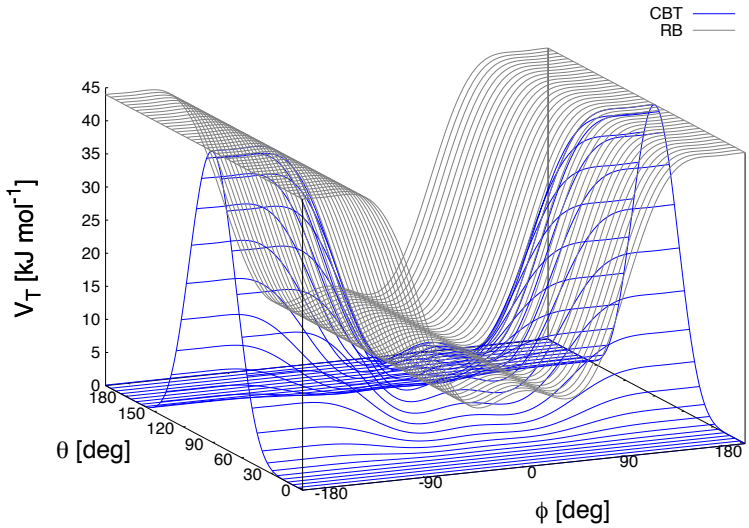

Figure 4.12: Blue: surface plot of the combined bending-torsion potential (4.68 with $k = 10$ kJ mol^{-1}, $a_0 = 2.41$, $a_1 = -2.95$, $a_2 = 0.36$, $a_3 = 1.33$) when, for simplicity, the bending angles behave the same ($\theta_1 = \theta_2 = \theta$). Gray: the same torsion potential without the $\sin^3\theta$ terms (Ryckaert-Bellemans type). ϕ is the dihedral angle.

$$V_a(\theta_{ijk}) = k\, f_n^a(\theta_{ijk}) \tag{4.71}$$
$$V_d(\phi_{ijkl}) = k\, f_n^d(\phi_{ijkl}) \tag{4.72}$$

where k is a force constant in units of energy and f is a cubic spline function; for details see 6.10.1. For each interaction, the force constant k and the table number n are specified in the topology. There are two different types of bonds, one that generates exclusions (type 8) and one that does not (type 9). For details see Table 5.5. The table files are supplied to the `mdrun` program. After the table file name an underscore, the letter "b" for bonds, "a" for angles or "d" for dihedrals and the table number must be appended. For example, a tabulated bond with $n = 0$ can be read from the file `table_b0.xvg`. Multiple tables can be supplied simply by adding files with different values of n, and are applied to the appropriate bonds, as specified in the topology (Table 5.5). The format for the table files is three fixed-format columns of any suitable width. These columns must contain x, $f(x)$, $-f'(x)$, and the values of x should be uniformly spaced. Requirements for entries in the topology are given in Table 5.5. The setup of the tables is as follows:
bonds: x is the distance in nm. For distances beyond the table length, `mdrun` will quit with an error message.
angles: x is the angle in degrees. The table should go from 0 up to and including 180 degrees; the derivative is taken in degrees.
dihedrals: x is the dihedral angle in degrees. The table should go from -180 up to and including 180 degrees; the IUPAC/IUB convention is used, *i.e.* zero is cis, the derivative is taken in degrees.

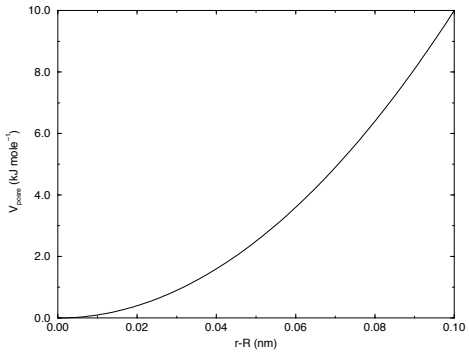

Figure 4.13: Position restraint potential.

4.3 Restraints

Special potentials are used for imposing restraints on the motion of the system, either to avoid disastrous deviations, or to include knowledge from experimental data. In either case they are not really part of the force field and the reliability of the parameters is not important. The potential forms, as implemented in GROMACS, are mentioned just for the sake of completeness. Restraints and constraints refer to quite different algorithms in GROMACS.

4.3.1 Position restraints

These are used to restrain particles to fixed reference positions $\boldsymbol{R}_i$. They can be used during equilibration in order to avoid drastic rearrangements of critical parts (*e.g.* to restrain motion in a protein that is subjected to large solvent forces when the solvent is not yet equilibrated). Another application is the restraining of particles in a shell around a region that is simulated in detail, while the shell is only approximated because it lacks proper interaction from missing particles outside the shell. Restraining will then maintain the integrity of the inner part. For spherical shells, it is a wise procedure to make the force constant depend on the radius, increasing from zero at the inner boundary to a large value at the outer boundary. This feature has not, however, been implemented in GROMACS.

The following form is used:

$$V_{pr}(\boldsymbol{r}_i) = \frac{1}{2} k_{pr} |\boldsymbol{r}_i - \boldsymbol{R}_i|^2 \tag{4.73}$$

The potential is plotted in Fig. 4.13.

The potential form can be rewritten without loss of generality as:

$$V_{pr}(\boldsymbol{r}_i) = \frac{1}{2} \left[k^x_{pr}(x_i - X_i)^2 \, \hat{\mathbf{x}} + k^y_{pr}(y_i - Y_i)^2 \, \hat{\mathbf{y}} + k^z_{pr}(z_i - Z_i)^2 \, \hat{\mathbf{z}} \right] \tag{4.74}$$

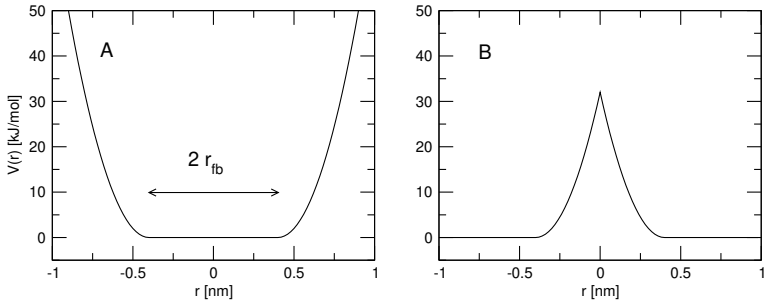

Figure 4.14: Flat-bottomed position restraint potential. (A) Not inverted, (B) inverted.

Now the forces are:

$$\begin{aligned} F_i^x &= -k_{pr}^x \, (x_i - X_i) \\ F_i^y &= -k_{pr}^y \, (y_i - Y_i) \\ F_i^z &= -k_{pr}^z \, (z_i - Z_i) \end{aligned} \tag{4.75}$$

Using three different force constants the position restraints can be turned on or off in each spatial dimension; this means that atoms can be harmonically restrained to a plane or a line. Position restraints are applied to a special fixed list of atoms. Such a list is usually generated by the `pdb2gmx` program.

4.3.2 Flat-bottomed position restraints

Flat-bottomed position restraints can be used to restrain particles to part of the simulation volume. No force acts on the restrained particle within the flat-bottomed region of the potential, however a harmonic force acts to move the particle to the flat-bottomed region if it is outside it. It is possible to apply normal and flat-bottomed position restraints on the same particle (however, only with the same reference position $\boldsymbol{R}_i$). The following general potential is used (Figure 4.14A):

$$V_{\mathrm{fb}}(\boldsymbol{r}_i) = \frac{1}{2} k_{\mathrm{fb}} [d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) - r_{\mathrm{fb}}]^2 \, H[d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) - r_{\mathrm{fb}}], \tag{4.76}$$

where $\boldsymbol{R}_i$ is the reference position, r_{fb} is the distance from the center with a flat potential, k_{fb} the force constant, and H is the Heaviside step function. The distance $d_g(\boldsymbol{r}_i; \boldsymbol{R}_i)$ from the reference position depends on the geometry g of the flat-bottomed potential.

The following geometries for the flat-bottomed potential are supported:
Sphere ($g = 1$): The particle is kept in a sphere of given radius. The force acts towards the center of the sphere. The following distance calculation is used:

$$d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) = |\boldsymbol{r}_i - \boldsymbol{R}_i| \tag{4.77}$$

Cylinder ($g = 6, 7, 8$): The particle is kept in a cylinder of given radius parallel to the x ($g = 6$), y ($g = 7$), or z-axis ($g = 8$). For backwards compatibility, setting $g = 2$ is mapped to $g = 8$ in the code so that old `.tpr` files and topologies work. The force from the flat-bottomed potential acts towards the axis of the cylinder. The component of the force parallel to the cylinder axis is zero. For a cylinder aligned along the z-axis:

$$d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) = \sqrt{(x_i - X_i)^2 + (y_i - Y_i)^2} \tag{4.78}$$

Layer ($g = 3, 4, 5$): The particle is kept in a layer defined by the thickness and the normal of the layer. The layer normal can be parallel to the x, y, or z-axis. The force acts parallel to the layer normal.

$$d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) = |x_i - X_i|, \quad \text{or} \quad d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) = |y_i - Y_i|, \quad \text{or} \quad d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) = |z_i - Z_i|. \tag{4.79}$$

It is possible to apply multiple independent flat-bottomed position restraints of different geometry on one particle. For example, applying a cylinder and a layer in z keeps a particle within a disk. Applying three layers in x, y, and z keeps the particle within a cuboid.

In addition, it is possible to invert the restrained region with the unrestrained region, leading to a potential that acts to keep the particle *outside* of the volume defined by $\boldsymbol{R}_i$, g, and r_{fb}. That feature is switched on by defining a negative r_{fb} in the topology. The following potential is used (Figure 4.14B):

$$V_{\mathrm{fb}}^{\mathrm{inv}}(\boldsymbol{r}_i) = \frac{1}{2} k_{\mathrm{fb}} [d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) - |r_{\mathrm{fb}}|]^2 \, H[-(d_g(\boldsymbol{r}_i; \boldsymbol{R}_i) - |r_{\mathrm{fb}}|)]. \tag{4.80}$$

4.3.3 Angle restraints

These are used to restrain the angle between two pairs of particles or between one pair of particles and the z-axis. The functional form is similar to that of a proper dihedral. For two pairs of atoms:

$$V_{ar}(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k, \boldsymbol{r}_l) = k_{ar}(1 - \cos(n(\theta - \theta_0))), \quad \text{where } \theta = \arccos\left(\frac{\boldsymbol{r}_j - \boldsymbol{r}_i}{\|\boldsymbol{r}_j - \boldsymbol{r}_i\|} \cdot \frac{\boldsymbol{r}_l - \boldsymbol{r}_k}{\|\boldsymbol{r}_l - \boldsymbol{r}_k\|} \right) \tag{4.81}$$

For one pair of atoms and the z-axis:

$$V_{ar}(\boldsymbol{r}_i, \boldsymbol{r}_j) = k_{ar}(1 - \cos(n(\theta - \theta_0))), \quad \text{where } \theta = \arccos\left(\frac{\boldsymbol{r}_j - \boldsymbol{r}_i}{\|\boldsymbol{r}_j - \boldsymbol{r}_i\|} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \tag{4.82}$$

A multiplicity (n) of 2 is useful when you do not want to distinguish between parallel and anti-parallel vectors. The equilibrium angle θ should be between 0 and 180 degrees for multiplicity 1 and between 0 and 90 degrees for multiplicity 2.

4.3.4 Dihedral restraints

These are used to restrain the dihedral angle ϕ defined by four particles as in an improper dihedral (sec. 4.2.12) but with a slightly modified potential. Using:

$$\phi' = (\phi - \phi_0) \ \mathrm{MOD}\ 2\pi \tag{4.83}$$

where ϕ_0 is the reference angle, the potential is defined as:

$$V_{dihr}(\phi') = \begin{cases} \frac{1}{2} k_{dihr}(\phi' - \phi_0 - \Delta\phi)^2 & \text{for } \phi' > \Delta\phi \\ 0 & \text{for } \phi' \leq \Delta\phi \end{cases} \tag{4.84}$$

where $\Delta\phi$ is a user defined angle and k_{dihr} is the force constant. **Note** that in the input in topology files, angles are given in degrees and force constants in kJ/mol/rad^2.

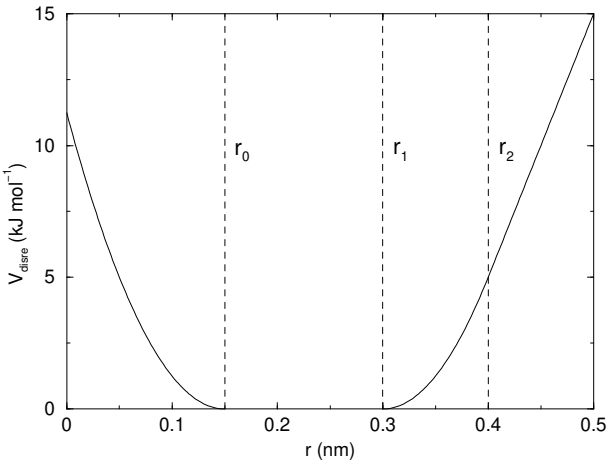

Figure 4.15: Distance Restraint potential.

4.3.5 Distance restraints

Distance restraints add a penalty to the potential when the distance between specified pairs of atoms exceeds a threshold value. They are normally used to impose experimental restraints from, for instance, experiments in nuclear magnetic resonance (NMR), on the motion of the system. Thus, MD can be used for structure refinement using NMR data. In GROMACS there are three ways to impose restraints on pairs of atoms:

- Simple harmonic restraints: use `[ bonds ]` type 6 . (see sec. 5.4.4).
- Piecewise linear/harmonic restraints: `[ bonds ]` type 10.
- Complex NMR distance restraints, optionally with pair, time and/or ensemble averaging.

The last two options will be detailed now.

The potential form for distance restraints is quadratic below a specified lower bound and between two specified upper bounds, and linear beyond the largest bound (see Fig. 4.15).

$$V_{dr}(r_{ij}) = \begin{cases} \frac{1}{2}k_{dr}(r_{ij}-r_0)^2 & \text{for } r_{ij} < r_0 \\ 0 & \text{for } r_0 \le r_{ij} < r_1 \\ \frac{1}{2}k_{dr}(r_{ij}-r_1)^2 & \text{for } r_1 \le r_{ij} < r_2 \\ \frac{1}{2}k_{dr}(r_2-r_1)(2r_{ij}-r_2-r_1) & \text{for } r_2 \le r_{ij} \end{cases} \tag{4.85}$$

The forces are

$$\boldsymbol{F}_i = \begin{cases} -k_{dr}(r_{ij}-r_0)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for } r_{ij} < r_0 \\ 0 & \text{for } r_0 \le r_{ij} < r_1 \\ -k_{dr}(r_{ij}-r_1)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for } r_1 \le r_{ij} < r_2 \\ -k_{dr}(r_2-r_1)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for } r_2 \le r_{ij} \end{cases} \tag{4.86}$$

For restraints not derived from NMR data, this functionality will usually suffice and a section of `[ bonds ]` type 10 can be used to apply individual restraints between pairs of atoms, see 5.8.1. For applying restraints derived from NMR measurements, more complex functionality might be required, which is provided through the `[ distance_restraints ]` section and is described below.

Time averaging

Distance restraints based on instantaneous distances can potentially reduce the fluctuations in a molecule significantly. This problem can be overcome by restraining to a *time averaged* distance [94]. The forces with time averaging are:

$$\boldsymbol{F}_i = \begin{cases} -k^a_{dr}(\bar{r}_{ij} - r_0)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for} \quad \bar{r}_{ij} < r_0 \\ 0 & \text{for} \quad r_0 \leq \bar{r}_{ij} < r_1 \\ -k^a_{dr}(\bar{r}_{ij} - r_1)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for} \quad r_1 \leq \bar{r}_{ij} < r_2 \\ -k^a_{dr}(r_2 - r_1)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for} \quad r_2 \leq \bar{r}_{ij} \end{cases} \tag{4.87}$$

where $\bar{r}_{ij}$ is given by an exponential running average with decay time τ:

$$\bar{r}_{ij} = < r_{ij}^{-3} >^{-1/3} \tag{4.88}$$

The force constant k^a_{dr} is switched on slowly to compensate for the lack of history at the beginning of the simulation:

$$k^a_{dr} = k_{dr}\left(1 - \exp\left(-\frac{t}{\tau}\right)\right) \tag{4.89}$$

Because of the time averaging, we can no longer speak of a distance restraint potential.

This way an atom can satisfy two incompatible distance restraints *on average* by moving between two positions. An example would be an amino acid side-chain that is rotating around its χ dihedral angle, thereby coming close to various other groups. Such a mobile side chain can give rise to multiple NOEs that can not be fulfilled by a single structure.

The computation of the time averaged distance in the `mdrun` program is done in the following fashion:

$$\begin{aligned} \overline{r^{-3}}_{ij}(0) &= r_{ij}(0)^{-3} \\ \overline{r^{-3}}_{ij}(t) &= \overline{r^{-3}}_{ij}(t - \Delta t)\, \exp\left(-\tfrac{\Delta t}{\tau}\right) + r_{ij}(t)^{-3}\left[1 - \exp\left(-\tfrac{\Delta t}{\tau}\right)\right] \end{aligned} \tag{4.90}$$

When a pair is within the bounds, it can still feel a force because the time averaged distance can still be beyond a bound. To prevent the protons from being pulled too close together, a mixed approach can be used. In this approach, the penalty is zero when the instantaneous distance is within the bounds, otherwise the violation is the square root of the product of the instantaneous violation and the time averaged violation:

$$\boldsymbol{F}_i = \begin{cases} k^a_{dr}\sqrt{(r_{ij} - r_0)(\bar{r}_{ij} - r_0)}\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for} \quad r_{ij} < r_0 \quad \text{and} \quad \bar{r}_{ij} < r_0 \\ -k^a_{dr}\min\left(\sqrt{(r_{ij} - r_1)(\bar{r}_{ij} - r_1)}, r_2 - r_1\right)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & \text{for} \quad r_{ij} > r_1 \quad \text{and} \quad \bar{r}_{ij} > r_1 \\ 0 & \text{otherwise} \end{cases} \tag{4.91}$$

Averaging over multiple pairs

Sometimes it is unclear from experimental data which atom pair gives rise to a single NOE, in other occasions it can be obvious that more than one pair contributes due to the symmetry of the system, *e.g.* a methyl group with three protons. For such a group, it is not possible to distinguish between the protons, therefore they should all be taken into account when calculating the distance between this methyl group and another proton (or group of protons). Due to the physical nature of magnetic resonance, the intensity of the NOE signal is inversely proportional to the sixth power of the inter-atomic distance. Thus, when combining atom pairs, a fixed list of N restraints may be taken together, where the apparent "distance" is given by:

$$r_N(t) = \left[\sum_{n=1}^{N} \bar{r}_n(t)^{-6} \right]^{-1/6} \tag{4.92}$$

where we use r_{ij} or eqn. 4.88 for the $\bar{r}_n$. The r_N of the instantaneous and time-averaged distances can be combined to do a mixed restraining, as indicated above. As more pairs of protons contribute to the same NOE signal, the intensity will increase, and the summed "distance" will be shorter than any of its components due to the reciprocal summation.

There are two options for distributing the forces over the atom pairs. In the conservative option, the force is defined as the derivative of the restraint potential with respect to the coordinates. This results in a conservative potential when time averaging is not used. The force distribution over the pairs is proportional to r^{-6}. This means that a close pair feels a much larger force than a distant pair, which might lead to a molecule that is "too rigid." The other option is an equal force distribution. In this case each pair feels $1/N$ of the derivative of the restraint potential with respect to r_N. The advantage of this method is that more conformations might be sampled, but the non-conservative nature of the forces can lead to local heating of the protons.

It is also possible to use *ensemble averaging* using multiple (protein) molecules. In this case the bounds should be lowered as in:

$$\begin{aligned} r_1 &= r_1 * M^{-1/6} \\ r_2 &= r_2 * M^{-1/6} \end{aligned} \tag{4.93}$$

where M is the number of molecules. The GROMACS preprocessor `grompp` can do this automatically when the appropriate option is given. The resulting "distance" is then used to calculate the scalar force according to:

$$\boldsymbol{F}_i = \begin{cases} 0 & r_N < r_1 \\ k_{dr}(r_N - r_1)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & r_1 \le r_N < r_2 \\ k_{dr}(r_2 - r_1)\frac{\boldsymbol{r}_{ij}}{r_{ij}} & r_N \ge r_2 \end{cases} \tag{4.94}$$

where i and j denote the atoms of all the pairs that contribute to the NOE signal.

Using distance restraints

A list of distance restrains based on NOE data can be added to a molecule definition in your topology file, like in the following example:

```
[ distance_restraints ]
; ai    aj    type    index    type'       low      up1      up2      fac
10      16       1        0        1       0.0      0.3      0.4      1.0
10      28       1        1        1       0.0      0.3      0.4      1.0
10      46       1        1        1       0.0      0.3      0.4      1.0
16      22       1        2        1       0.0      0.3      0.4      2.5
16      34       1        3        1       0.0      0.5      0.6      1.0
```

In this example a number of features can be found. In columns `ai` and `aj` you find the atom numbers of the particles to be restrained. The `type` column should always be 1. As explained in 4.3.5, multiple distances can contribute to a single NOE signal. In the topology this can be set using the `index` column. In our example, the restraints 10-28 and 10-46 both have index 1, therefore they are treated simultaneously. An extra requirement for treating restraints together is that the restraints must be on successive lines, without any other intervening restraint. The `type'` column will usually be 1, but can be set to 2 to obtain a distance restraint that will never be time- and ensemble-averaged; this can be useful for restraining hydrogen bonds. The columns `low`, `up1`, and `up2` hold the values of r_0, r_1, and r_2 from eqn. 4.85. In some cases it can be useful to have different force constants for some restraints; this is controlled by the column `fac`. The force constant in the parameter file is multiplied by the value in the column `fac` for each restraint.

4.3.6 Orientation restraints

This section describes how orientations between vectors, as measured in certain NMR experiments, can be calculated and restrained in MD simulations. The presented refinement methodology and a comparison of results with and without time and ensemble averaging have been published [95].

Theory

In an NMR experiment, orientations of vectors can be measured when a molecule does not tumble completely isotropically in the solvent. Two examples of such orientation measurements are residual dipolar couplings (between two nuclei) or chemical shift anisotropies. An observable for a vector $\boldsymbol{r}_i$ can be written as follows:

$$\delta_i = \frac{2}{3}\mathrm{tr}(\mathbf{S}\mathbf{D}_i) \tag{4.95}$$

where $\mathbf{S}$ is the dimensionless order tensor of the molecule. The tensor $\mathbf{D}_i$ is given by:

$$\mathbf{D}_i = \frac{c_i}{\|\boldsymbol{r}_i\|^\alpha}\begin{pmatrix} 3xx-1 & 3xy & 3xz \\ 3xy & 3yy-1 & 3yz \\ 3xz & 3yz & 3zz-1 \end{pmatrix} \tag{4.96}$$

$$\text{with:} \quad x = \frac{r_{i,x}}{\|\boldsymbol{r}_i\|}, \quad y = \frac{r_{i,y}}{\|\boldsymbol{r}_i\|}, \quad z = \frac{r_{i,z}}{\|\boldsymbol{r}_i\|} \tag{4.97}$$

For a dipolar coupling $\boldsymbol{r}_i$ is the vector connecting the two nuclei, $\alpha = 3$ and the constant c_i is given by:

$$c_i = \frac{\mu_0}{4\pi}\gamma_1^i\gamma_2^i\frac{\hbar}{4\pi} \tag{4.98}$$

where γ_1^i and γ_2^i are the gyromagnetic ratios of the two nuclei.

The order tensor is symmetric and has trace zero. Using a rotation matrix $\mathbf{T}$ it can be transformed into the following form:

$$\mathbf{T}^T\mathbf{S}\mathbf{T} = s\begin{pmatrix} -\frac{1}{2}(1-\eta) & 0 & 0 \\ 0 & -\frac{1}{2}(1+\eta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.99}$$

where $-1 \leq s \leq 1$ and $0 \leq \eta \leq 1$. s is called the order parameter and η the asymmetry of the order tensor $\mathbf{S}$. When the molecule tumbles isotropically in the solvent, s is zero, and no orientational effects can be observed because all δ_i are zero.

Calculating orientations in a simulation

For reasons which are explained below, the $\mathbf{D}$ matrices are calculated which respect to a reference orientation of the molecule. The orientation is defined by a rotation matrix $\mathbf{R}$, which is needed to least-squares fit the current coordinates of a selected set of atoms onto a reference conformation. The reference conformation is the starting conformation of the simulation. In case of ensemble averaging, which will be treated later, the structure is taken from the first subsystem. The calculated $\mathbf{D}_i^c$ matrix is given by:

$$\mathbf{D}_i^c(t) = \mathbf{R}(t)\mathbf{D}_i(t)\mathbf{R}^T(t) \tag{4.100}$$

The calculated orientation for vector i is given by:

$$\delta_i^c(t) = \frac{2}{3}\mathrm{tr}(\mathbf{S}(t)\mathbf{D}_i^c(t)) \tag{4.101}$$

The order tensor $\mathbf{S}(t)$ is usually unknown. A reasonable choice for the order tensor is the tensor which minimizes the (weighted) mean square difference between the calculated and the observed orientations:

$$MSD(t) = \left(\sum_{i=1}^{N} w_i\right)^{-1}\sum_{i=1}^{N} w_i(\delta_i^c(t) - \delta_i^{exp})^2 \tag{4.102}$$

To properly combine different types of measurements, the unit of w_i should be such that all terms are dimensionless. This means the unit of w_i is the unit of δ_i to the power -2. **Note** that scaling all w_i with a constant factor does not influence the order tensor.

Time averaging

Since the tensors $\mathbf{D}_i$ fluctuate rapidly in time, much faster than can be observed in an experiment, they should be averaged over time in the simulation. However, in a simulation the time and the number of copies of a molecule are limited. Usually one can not obtain a converged average of the $\mathbf{D}_i$ tensors over all orientations of the molecule. If one assumes that the average orientations of

the $\boldsymbol{r}_i$ vectors within the molecule converge much faster than the tumbling time of the molecule, the tensor can be averaged in an axis system that rotates with the molecule, as expressed by equation (4.100). The time-averaged tensors are calculated using an exponentially decaying memory function:

$$\mathbf{D}_i^a(t) = \frac{\int_{u=t_0}^{t} \mathbf{D}_i^c(u) \exp\left(-\frac{t-u}{\tau}\right) \mathrm{d}u}{\int_{u=t_0}^{t} \exp\left(-\frac{t-u}{\tau}\right) \mathrm{d}u} \tag{4.103}$$

Assuming that the order tensor $\mathbf{S}$ fluctuates slower than the $\mathbf{D}_i$, the time-averaged orientation can be calculated as:

$$\delta_i^a(t) = \frac{2}{3}\mathrm{tr}(\mathbf{S}(t)\mathbf{D}_i^a(t)) \tag{4.104}$$

where the order tensor $\mathbf{S}(t)$ is calculated using expression (4.102) with $\delta_i^c(t)$ replaced by $\delta_i^a(t)$.

Restraining

The simulated structure can be restrained by applying a force proportional to the difference between the calculated and the experimental orientations. When no time averaging is applied, a proper potential can be defined as:

$$V = \frac{1}{2}k\sum_{i=1}^{N} w_i(\delta_i^c(t) - \delta_i^{exp})^2 \tag{4.105}$$

where the unit of k is the unit of energy. Thus the effective force constant for restraint i is kw_i. The forces are given by minus the gradient of V. The force $\boldsymbol{F}_i$ working on vector $\boldsymbol{r}_i$ is:

$$\begin{aligned}
\boldsymbol{F}_i(t) &= -\frac{\mathrm{d}V}{\mathrm{d}\boldsymbol{r}_i} \\
&= -kw_i(\delta_i^c(t) - \delta_i^{exp})\frac{\mathrm{d}\delta_i(t)}{\mathrm{d}\boldsymbol{r}_i} \\
&= -kw_i(\delta_i^c(t) - \delta_i^{exp})\frac{2c_i}{\|\boldsymbol{r}\|^{2+\alpha}}\left(2\mathbf{R}^T\mathbf{S}\mathbf{R}\boldsymbol{r}_i - \frac{2+\alpha}{\|\boldsymbol{r}\|^2}\mathrm{tr}(\mathbf{R}^T\mathbf{S}\mathbf{R}\boldsymbol{r}_i\boldsymbol{r}_i^T)\boldsymbol{r}_i\right)
\end{aligned}$$

Ensemble averaging

Ensemble averaging can be applied by simulating a system of M subsystems that each contain an identical set of orientation restraints. The systems only interact via the orientation restraint potential which is defined as:

$$V = M\frac{1}{2}k\sum_{i=1}^{N} w_i\langle\delta_i^c(t) - \delta_i^{exp}\rangle^2 \tag{4.106}$$

The force on vector $\boldsymbol{r}_{i,m}$ in subsystem m is given by:

$$\boldsymbol{F}_{i,m}(t) = -\frac{\mathrm{d}V}{\mathrm{d}\boldsymbol{r}_{i,m}} = -kw_i\langle\delta_i^c(t) - \delta_i^{exp}\rangle\frac{\mathrm{d}\delta_{i,m}^c(t)}{\mathrm{d}\boldsymbol{r}_{i,m}} \tag{4.107}$$

Time averaging

When using time averaging it is not possible to define a potential. We can still define a quantity that gives a rough idea of the energy stored in the restraints:

$$V = M\frac{1}{2}k^a \sum_{i=1}^{N} w_i \langle \delta_i^a(t) - \delta_i^{exp} \rangle^2 \tag{4.108}$$

The force constant k_a is switched on slowly to compensate for the lack of history at times close to t_0. It is exactly proportional to the amount of average that has been accumulated:

$$k^a = k\,\frac{1}{\tau}\int_{u=t_0}^{t} \exp\left(-\frac{t-u}{\tau}\right) \mathrm{d}u \tag{4.109}$$

What really matters is the definition of the force. It is chosen to be proportional to the square root of the product of the time-averaged and the instantaneous deviation. Using only the time-averaged deviation induces large oscillations. The force is given by:

$$\boldsymbol{F}_{i,m}(t) = \begin{cases} 0 & \text{for} \quad a\,b \le 0 \\ k^a w_i \dfrac{a}{|a|}\sqrt{a\,b}\,\dfrac{\mathrm{d}\delta_{i,m}^c(t)}{\mathrm{d}\boldsymbol{r}_{i,m}} & \text{for} \quad a\,b > 0 \end{cases} \tag{4.110}$$

$$\begin{aligned} a &= \langle \delta_i^a(t) - \delta_i^{exp} \rangle \\ b &= \langle \delta_i^c(t) - \delta_i^{exp} \rangle \end{aligned}$$

Using orientation restraints

Orientation restraints can be added to a molecule definition in the topology file in the section `[ orientation_restraints ]`. Here we give an example section containing five N-H residual dipolar coupling restraints:

```
[ orientation_restraints ]
; ai   aj  type  exp.  label  alpha    const.     obs.   weight
;                             Hz       nm^3       Hz    Hz^-2
  31   32     1     1      3      3     6.083    -6.73     1.0
  43   44     1     1      4      3     6.083    -7.87     1.0
  55   56     1     1      5      3     6.083    -7.13     1.0
  65   66     1     1      6      3     6.083    -2.57     1.0
  73   74     1     1      7      3     6.083    -2.10     1.0
```

The unit of the observable is Hz, but one can choose any other unit. In columns `ai` and `aj` you find the atom numbers of the particles to be restrained. The `type` column should always be 1. The `exp.` column denotes the experiment number, starting at 1. For each experiment a separate order tensor $\mathbf{S}$ is optimized. The label should be a unique number larger than zero for each restraint. The `alpha` column contains the power α that is used in equation (4.96) to calculate the orientation. The `const.` column contains the constant c_i used in the same equation. The constant should

have the unit of the observable times nm$^{\alpha}$. The column `obs.` contains the observable, in any unit you like. The last column contains the weights w_i; the unit should be the inverse of the square of the unit of the observable.

Some parameters for orientation restraints can be specified in the `grompp.mdp` file, for a study of the effect of different force constants and averaging times and ensemble averaging see [95].

4.4 Polarization

Polarization can be treated by GROMACS by attaching shell (Drude) particles to atoms and/or virtual sites. The energy of the shell particle is then minimized at each time step in order to remain on the Born-Oppenheimer surface.

4.4.1 Simple polarization

This is merely a harmonic potential with equilibrium distance 0.

4.4.2 Water polarization

A special potential for water that allows anisotropic polarization of a single shell particle [44].

4.4.3 Thole polarization

Based on early work by Thole [96], Roux and coworkers have implemented potentials for molecules like ethanol [97, 98, 99]. Within such molecules, there are intra-molecular interactions between shell particles, however these must be screened because full Coulomb would be too strong. The potential between two shell particles i and j is:

$$V_{thole} = \frac{q_i q_j}{r_{ij}} \left[1 - \left(1 + \frac{\bar{r}_{ij}}{2} \right) \exp^{-\bar{r}_{ij}} \right] \tag{4.111}$$

Note that there is a sign error in Equation 1 of Noskov *et al.* [99]:

$$\bar{r}_{ij} = a \frac{r_{ij}}{(\alpha_i \alpha_j)^{1/6}} \tag{4.112}$$

where a is a magic (dimensionless) constant, usually chosen to be 2.6 [99]; α_i and α_j are the polarizabilities of the respective shell particles.

4.5 Free energy interactions

This section describes the λ-dependence of the potentials used for free energy calculations (see sec. 3.12). All common types of potentials and constraints can be interpolated smoothly from state A ($\lambda = 0$) to state B ($\lambda = 1$) and vice versa. All bonded interactions are interpolated by linear

interpolation of the interaction parameters. Non-bonded interactions can be interpolated linearly or via soft-core interactions.

Starting in GROMACS 4.6, λ is a vector, allowing different components of the free energy transformation to be carried out at different rates. Coulomb, Lennard-Jones, bonded, and restraint terms can all be controlled independently, as described in the `.mdp` options.

Harmonic potentials

The example given here is for the bond potential, which is harmonic in GROMACS. However, these equations apply to the angle potential and the improper dihedral potential as well.

$$V_b = \frac{1}{2}\left[(1-\lambda)k_b^A + \lambda k_b^B\right]\left[b-(1-\lambda)b_0^A - \lambda b_0^B\right]^2 \tag{4.113}$$

$$\frac{\partial V_b}{\partial \lambda} = \frac{1}{2}(k_b^B - k_b^A)\left[b-(1-\lambda)b_0^A + \lambda b_0^B\right]^2 + (b_0^A - b_0^B)\left[b-(1-\lambda)b_0^A - \lambda b_0^B\right]\left[(1-\lambda)k_b^A + \lambda k_b^B\right] \tag{4.114}$$

GROMOS-96 bonds and angles

Fourth-power bond stretching and cosine-based angle potentials are interpolated by linear interpolation of the force constant and the equilibrium position. Formulas are not given here.

Proper dihedrals

For the proper dihedrals, the equations are somewhat more complicated:

$$V_d = \left[(1-\lambda)k_d^A + \lambda k_d^B\right]\left(1+\cos\left[n_\phi\phi - (1-\lambda)\phi_s^A - \lambda\phi_s^B\right]\right) \tag{4.115}$$

$$\frac{\partial V_d}{\partial \lambda} = (k_d^B - k_d^A)\left(1+\cos\left[n_\phi\phi - (1-\lambda)\phi_s^A - \lambda\phi_s^B\right]\right) + (\phi_s^B - \phi_s^A)\left[(1-\lambda)k_d^A - \lambda k_d^B\right]\sin\left[n_\phi\phi - (1-\lambda)\phi_s^A - \lambda\phi_s^B\right] \tag{4.116}$$

Note: that the multiplicity n_ϕ can not be parameterized because the function should remain periodic on the interval $[0, 2\pi]$.

Tabulated bonded interactions

For tabulated bonded interactions only the force constant can interpolated:

$$V = ((1-\lambda)k^A + \lambda k^B)\, f \tag{4.117}$$

$$\frac{\partial V}{\partial \lambda} = (k^B - k^A)\, f \tag{4.118}$$

Coulomb interaction

The Coulomb interaction between two particles of which the charge varies with λ is:

$$V_c = \frac{f}{\varepsilon_{rf} r_{ij}} \left[(1-\lambda) q_i^A q_j^A + \lambda\, q_i^B q_j^B \right] \tag{4.119}$$

$$\frac{\partial V_c}{\partial \lambda} = \frac{f}{\varepsilon_{rf} r_{ij}} \left[-q_i^A q_j^A + q_i^B q_j^B \right] \tag{4.120}$$

where $f = \frac{1}{4\pi\varepsilon_0} = 138.935\,458$ (see chapter 2).

Coulomb interaction with reaction field

The Coulomb interaction including a reaction field, between two particles of which the charge varies with λ is:

$$V_c = f \left[\frac{1}{r_{ij}} + k_{rf}\, r_{ij}^2 - c_{rf} \right] \left[(1-\lambda) q_i^A q_j^A + \lambda\, q_i^B q_j^B \right] \tag{4.121}$$

$$\frac{\partial V_c}{\partial \lambda} = f \left[\frac{1}{r_{ij}} + k_{rf}\, r_{ij}^2 - c_{rf} \right] \left[-q_i^A q_j^A + q_i^B q_j^B \right] \tag{4.122}$$

Note that the constants k_{rf} and c_{rf} are defined using the dielectric constant ε_{rf} of the medium (see sec. 4.1.4).

Lennard-Jones interaction

For the Lennard-Jones interaction between two particles of which the *atom type* varies with λ we can write:

$$V_{LJ} = \frac{(1-\lambda) C_{12}^A + \lambda\, C_{12}^B}{r_{ij}^{12}} - \frac{(1-\lambda) C_6^A + \lambda\, C_6^B}{r_{ij}^6} \tag{4.123}$$

$$\frac{\partial V_{LJ}}{\partial \lambda} = \frac{C_{12}^B - C_{12}^A}{r_{ij}^{12}} - \frac{C_6^B - C_6^A}{r_{ij}^6} \tag{4.124}$$

It should be noted that it is also possible to express a pathway from state A to state B using σ and ϵ (see eqn. 4.5). It may seem to make sense physically to vary the force field parameters σ and ϵ rather than the derived parameters C_{12} and C_6. However, the difference between the pathways in parameter space is not large, and the free energy itself does not depend on the pathway, so we use the simple formulation presented above.

Kinetic Energy

When the mass of a particle changes, there is also a contribution of the kinetic energy to the free energy (note that we can not write the momentum $\boldsymbol{p}$ as m$\boldsymbol{v}$, since that would result in the sign of $\frac{\partial E_k}{\partial \lambda}$ being incorrect [100]):

$$E_k = \frac{1}{2}\frac{\boldsymbol{p}^2}{(1-\lambda)m^A + \lambda m^B} \tag{4.125}$$

$$\frac{\partial E_k}{\partial \lambda} = -\frac{1}{2}\frac{\boldsymbol{p}^2(m^B - m^A)}{((1-\lambda)m^A + \lambda m^B)^2} \tag{4.126}$$

after taking the derivative, we *can* insert $\boldsymbol{p}$ = m$\boldsymbol{v}$, such that:

$$\frac{\partial E_k}{\partial \lambda} = -\frac{1}{2}\boldsymbol{v}^2(m^B - m^A) \tag{4.127}$$

Constraints

The constraints are formally part of the Hamiltonian, and therefore they give a contribution to the free energy. In GROMACS this can be calculated using the LINCS or the SHAKE algorithm. If we have $k = 1 \ldots K$ constraint equations g_k for LINCS, then

$$g_k = |\boldsymbol{r}_k| - d_k \tag{4.128}$$

where $\boldsymbol{r}_k$ is the displacement vector between two particles and d_k is the constraint distance between the two particles. We can express the fact that the constraint distance has a λ dependency by

$$d_k = (1-\lambda)d_k^A + \lambda d_k^B \tag{4.129}$$

Thus the λ-dependent constraint equation is

$$g_k = |\boldsymbol{r}_k| - \left((1-\lambda)d_k^A + \lambda d_k^B\right). \tag{4.130}$$

The (zero) contribution G to the Hamiltonian from the constraints (using Lagrange multipliers λ_k, which are logically distinct from the free-energy λ) is

$$G = \sum_k^K \lambda_k g_k \tag{4.131}$$

$$\frac{\partial G}{\partial \lambda} = \frac{\partial G}{\partial d_k}\frac{\partial d_k}{\partial \lambda} \tag{4.132}$$

$$= -\sum_k^K \lambda_k \left(d_k^B - d_k^A\right) \tag{4.133}$$

For SHAKE, the constraint equations are

$$g_k = \boldsymbol{r}_k^2 - d_k^2 \tag{4.134}$$

with d_k as before, so

$$\frac{\partial G}{\partial \lambda} = -2\sum_k^K \lambda_k \left(d_k^B - d_k^A\right) \tag{4.135}$$

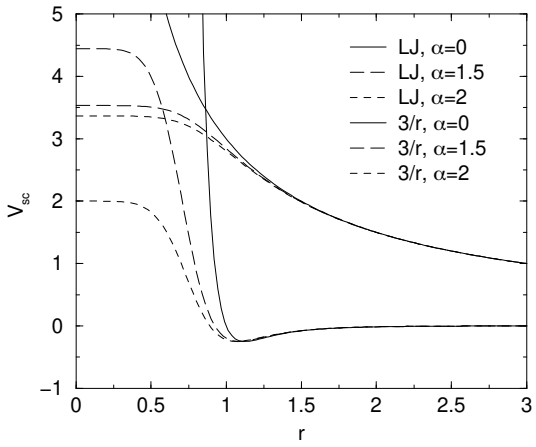

Figure 4.16: Soft-core interactions at $\lambda = 0.5$, with $p = 2$ and $C_6^A = C_{12}^A = C_6^B = C_{12}^B = 1$.

4.5.1 Soft-core interactions

In a free-energy calculation where particles grow out of nothing, or particles disappear, using the the simple linear interpolation of the Lennard-Jones and Coulomb potentials as described in Equations 4.124 and 4.122 may lead to poor convergence. When the particles have nearly disappeared, or are close to appearing (at λ close to 0 or 1), the interaction energy will be weak enough for particles to get very close to each other, leading to large fluctuations in the measured values of $\partial V/\partial\lambda$ (which, because of the simple linear interpolation, depends on the potentials at both the endpoints of λ).

To circumvent these problems, the singularities in the potentials need to be removed. This can be done by modifying the regular Lennard-Jones and Coulomb potentials with "soft-core" potentials that limit the energies and forces involved at λ values between 0 and 1, but not *at* $\lambda = 0$ or 1.

In GROMACS the soft-core potentials V_{sc} are shifted versions of the regular potentials, so that the singularity in the potential and its derivatives at $r = 0$ is never reached:

$$V_{sc}(r) = (1-\lambda)V^A(r_A) + \lambda V^B(r_B) \tag{4.136}$$

$$r_A = \left(\alpha\sigma_A^6\lambda^p + r^6\right)^{\frac{1}{6}} \tag{4.137}$$

$$r_B = \left(\alpha\sigma_B^6(1-\lambda)^p + r^6\right)^{\frac{1}{6}} \tag{4.138}$$

where V^A and V^B are the normal "hard core" Van der Waals or electrostatic potentials in state A ($\lambda = 0$) and state B ($\lambda = 1$) respectively, α is the soft-core parameter (set with `sc_alpha` in the `.mdp` file), p is the soft-core λ power (set with `sc_power`), σ is the radius of the interaction, which is $(C_{12}/C_6)^{1/6}$ or an input parameter (`sc_sigma`) when C_6 or C_{12} is zero.

For intermediate λ, r_A and r_B alter the interactions very little for $r > \alpha^{1/6}\sigma$ and quickly switch the soft-core interaction to an almost constant value for smaller r (Fig. 4.16). The force is:

$$F_{sc}(r) = -\frac{\partial V_{sc}(r)}{\partial r} = (1-\lambda)F^A(r_A)\left(\frac{r}{r_A}\right)^5 + \lambda F^B(r_B)\left(\frac{r}{r_B}\right)^5 \tag{4.139}$$

where F^A and F^B are the "hard core" forces. The contribution to the derivative of the free energy is:

$$\begin{aligned}\frac{\partial V_{sc}(r)}{\partial\lambda} &= V^B(r_B) - V^A(r_A) + (1-\lambda)\frac{\partial V^A(r_A)}{\partial r_A}\frac{\partial r_A}{\partial\lambda} + \lambda\frac{\partial V^B(r_B)}{\partial r_B}\frac{\partial r_B}{\partial\lambda} \\ &= V^B(r_B) - V^A(r_A) + \\ &\quad \frac{p\alpha}{6}\left[\lambda F^B(r_B) r_B^{-5}\sigma_B^6(1-\lambda)^{p-1} - (1-\lambda)F^A(r_A)r_A^{-5}\sigma_A^6\lambda^{p-1}\right] \quad (4.140)\end{aligned}$$

The original GROMOS Lennard-Jones soft-core function [101] uses $p = 2$, but $p = 1$ gives a smoother $\partial H/\partial\lambda$ curve.

Another issue that should be considered is the soft-core effect of hydrogens without Lennard-Jones interaction. Their soft-core σ is set with `sc-sigma` in the `.mdp` file. These hydrogens produce peaks in $\partial H/\partial\lambda$ at λ is 0 and/or 1 for $p = 1$ and close to 0 and/or 1 with $p = 2$. Lowering `sc-sigma` will decrease this effect, but it will also increase the interactions with hydrogens relative to the other interactions in the soft-core state.

When soft-core potentials are selected (by setting `sc-alpha` >0), and the Coulomb and Lennard-Jones potentials are turned on or off sequentially, then the Coulombic interaction is turned off linearly, rather than using soft-core interactions, which should be less statistically noisy in most cases. This behavior can be overwritten by using the mdp option `sc-coul` to `yes`. Note that the `sc-coul` is only taken into account when lambda states are used, not with `couple-lambda0` / `couple-lambda1`, and you can still turn off soft-core interactions by setting `sc-alpha=0`. Additionally, the soft-core interaction potential is only applied when either the A or B state has zero interaction potential. If both A and B states have nonzero interaction potential, default linear scaling described above is used. When both Coulombic and Lennard-Jones interactions are turned off simultaneously, a soft-core potential is used, and a hydrogen is being introduced or deleted, the sigma is set to `sc-sigma-min`, which itself defaults to `sc-sigma-default`.

Recently, a new formulation of the soft-core approach has been derived that in most cases gives lower and more even statistical variance than the standard soft-core path described above. [102, 103] Specifically, we have:

$$V_{sc}(r) = (1-\lambda)V^A(r_A) + \lambda V^B(r_B) \quad (4.141)$$

$$r_A = \left(\alpha\sigma_A^{48}\lambda^p + r^{48}\right)^{\frac{1}{48}} \quad (4.142)$$

$$r_B = \left(\alpha\sigma_B^{48}(1-\lambda)^p + r^{48}\right)^{\frac{1}{48}} \quad (4.143)$$

This "1-1-48" path is also implemented in GROMACS. Note that for this path the soft core α should satisfy $0.001 < \alpha < 0.003$, rather than $\alpha \approx 0.5$.

4.6 Methods

4.6.1 Exclusions and 1-4 Interactions.

Atoms within a molecule that are close by in the chain, *i.e.* atoms that are covalently bonded, or linked by one or two atoms are called *first neighbors, second neighbors* and *third neighbors*,

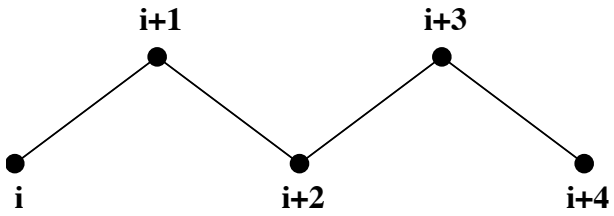

Figure 4.17: Atoms along an alkane chain.

respectively (see Fig. 4.17). Since the interactions of atom **i** with atoms **i+1** and **i+2** are mainly quantum mechanical, they can not be modeled by a Lennard-Jones potential. Instead it is assumed that these interactions are adequately modeled by a harmonic bond term or constraint (**i, i+1**) and a harmonic angle term (**i, i+2**). The first and second neighbors (atoms **i+1** and **i+2**) are therefore *excluded* from the Lennard-Jones interaction list of atom **i**; atoms **i+1** and **i+2** are called *exclusions* of atom **i**.

For third neighbors, the normal Lennard-Jones repulsion is sometimes still too strong, which means that when applied to a molecule, the molecule would deform or break due to the internal strain. This is especially the case for carbon-carbon interactions in a *cis*-conformation (*e.g. cis*-butane). Therefore, for some of these interactions, the Lennard-Jones repulsion has been reduced in the GROMOS force field, which is implemented by keeping a separate list of 1-4 and normal Lennard-Jones parameters. In other force fields, such as OPLS [104], the standard Lennard-Jones parameters are reduced by a factor of two, but in that case also the dispersion (r^{-6}) and the Coulomb interaction are scaled. GROMACS can use either of these methods.

4.6.2 Charge Groups

In principle, the force calculation in MD is an $O(N^2)$ problem. Therefore, we apply a cut-off for non-bonded force (NBF) calculations; only the particles within a certain distance of each other are interacting. This reduces the cost to $O(N)$ (typically $100N$ to $200N$) of the NBF. It also introduces an error, which is, in most cases, acceptable, except when applying the cut-off implies the creation of charges, in which case you should consider using the lattice sum methods provided by GROMACS.

Consider a water molecule interacting with another atom. If we would apply a plain cut-off on an atom-atom basis we might include the atom-oxygen interaction (with a charge of -0.82) without the compensating charge of the protons, and as a result, induce a large dipole moment over the system. Therefore, we have to keep groups of atoms with total charge 0 together. These groups are called *charge groups*. Note that with a proper treatment of long-range electrostatics (e.g. particle-mesh Ewald (sec. 4.8.2), keeping charge groups together is not required.

4.6.3 Treatment of Cut-offs in the group scheme

GROMACS is quite flexible in treating cut-offs, which implies there can be quite a number of parameters to set. These parameters are set in the input file for `grompp`. There are two sort of parameters that affect the cut-off interactions; you can select which type of interaction to use in

each case, and which cut-offs should be used in the neighbor searching.

For both Coulomb and van der Waals interactions there are interaction type selectors (termed `vdwtype` and `coulombtype`) and two parameters, for a total of six non-bonded interaction parameters. See the User Guide for a complete description of these parameters.

In the group cut-off scheme, all of the interaction functions in Table 4.2 require that neighbor searching be done with a radius at least as large as the r_c specified for the functional form, because of the use of charge groups. The extra radius is typically of the order of 0.25 nm (roughly the largest distance between two atoms in a charge group plus the distance a charge group can diffuse within neighbor list updates).

Type		Parameters
Coulomb	Plain cut-off	r_c, ε_r
	Reaction field	r_c, ε_{rf}
	Shift function	r_1, r_c, ε_r
	Switch function	r_1, r_c, ε_r
VdW	Plain cut-off	r_c
	Shift function	r_1, r_c
	Switch function	r_1, r_c

Table 4.2: Parameters for the different functional forms of the non-bonded interactions.

4.7 Virtual interaction sites

Virtual interaction sites (called dummy atoms in GROMACS versions before 3.3) can be used in GROMACS in a number of ways. We write the position of the virtual site $\boldsymbol{r}_s$ as a function of the positions of other particles $\boldsymbol{r}_i$: $\boldsymbol{r}_s = f(\boldsymbol{r}_1..\boldsymbol{r}_n)$. The virtual site, which may carry charge or be involved in other interactions, can now be used in the force calculation. The force acting on the virtual site must be redistributed over the particles with mass in a consistent way. A good way to do this can be found in ref. [105]. We can write the potential energy as:

$$V = V(\boldsymbol{r}_s, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_n) = V^*(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n) \tag{4.144}$$

The force on the particle i is then:

$$\boldsymbol{F}_i = -\frac{\partial V^*}{\partial \boldsymbol{r}_i} = -\frac{\partial V}{\partial \boldsymbol{r}_i} - \frac{\partial V}{\partial \boldsymbol{r}_s}\frac{\partial \boldsymbol{r}_s}{\partial \boldsymbol{r}_i} = \boldsymbol{F}_i^{direct} + \boldsymbol{F}'_i \tag{4.145}$$

The first term is the normal force. The second term is the force on particle i due to the virtual site, which can be written in tensor notation:

$$\boldsymbol{F}'_i = \begin{bmatrix} \frac{\partial x_s}{\partial x_i} & \frac{\partial y_s}{\partial x_i} & \frac{\partial z_s}{\partial x_i} \\ \frac{\partial x_s}{\partial y_i} & \frac{\partial y_s}{\partial y_i} & \frac{\partial z_s}{\partial y_i} \\ \frac{\partial x_s}{\partial z_i} & \frac{\partial y_s}{\partial z_i} & \frac{\partial z_s}{\partial z_i} \end{bmatrix} \boldsymbol{F}_s \tag{4.146}$$

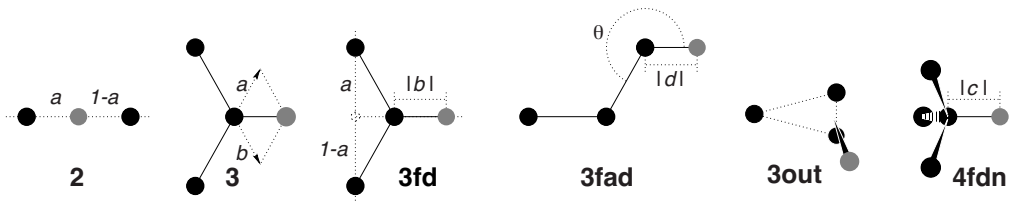

Figure 4.18: The six different types of virtual site construction in GROMACS. The constructing atoms are shown as black circles, the virtual sites in gray.

where $\boldsymbol{F}_s$ is the force on the virtual site and x_s, y_s and z_s are the coordinates of the virtual site. In this way, the total force and the total torque are conserved [105].

The computation of the virial (eqn. 3.24) is non-trivial when virtual sites are used. Since the virial involves a summation over all the atoms (rather than virtual sites), the forces must be redistributed from the virtual sites to the atoms (using eqn. 4.146) *before* computation of the virial. In some special cases where the forces on the atoms can be written as a linear combination of the forces on the virtual sites (types 2 and 3 below) there is no difference between computing the virial before and after the redistribution of forces. However, in the general case redistribution should be done first.

There are six ways to construct virtual sites from surrounding atoms in GROMACS, which we classify by the number of constructing atoms. **Note** that all site types mentioned can be constructed from types 3fd (normalized, in-plane) and 3out (non-normalized, out of plane). However, the amount of computation involved increases sharply along this list, so we strongly recommended using the first adequate virtual site type that will be sufficient for a certain purpose. Fig. 4.18 depicts 6 of the available virtual site constructions. The conceptually simplest construction types are linear combinations:

$$\boldsymbol{r}_s = \sum_{i=1}^{N} w_i \, \boldsymbol{r}_i \tag{4.147}$$

The force is then redistributed using the same weights:

$$\boldsymbol{F}'_i = w_i \, \boldsymbol{F}_s \tag{4.148}$$

The types of virtual sites supported in GROMACS are given in the list below. Constructing atoms in virtual sites can be virtual sites themselves, but only if they are higher in the list, i.e. virtual sites can be constructed from "particles" that are simpler virtual sites.

2. As a linear combination of two atoms (Fig. 4.18 2):

$$w_i = 1 - a \, , \quad w_j = a \tag{4.149}$$

   In this case the virtual site is on the line through atoms i and j.

3. As a linear combination of three atoms (Fig. 4.18 3):

$$w_i = 1 - a - b \, , \quad w_j = a \, , \quad w_k = b \tag{4.150}$$

   In this case the virtual site is in the plane of the other three particles.

3fd. In the plane of three atoms, with a fixed distance (Fig. 4.18 3fd):

$$\boldsymbol{r}_s = \boldsymbol{r}_i + b\frac{\boldsymbol{r}_{ij} + a\boldsymbol{r}_{jk}}{|\boldsymbol{r}_{ij} + a\boldsymbol{r}_{jk}|} \tag{4.151}$$

In this case the virtual site is in the plane of the other three particles at a distance of $|b|$ from i. The force on particles i, j and k due to the force on the virtual site can be computed as:

$$\begin{aligned} \boldsymbol{F}'_i &= \boldsymbol{F}_s - \gamma(\boldsymbol{F}_s - \boldsymbol{p}) \\ \boldsymbol{F}'_j &= (1-a)\gamma(\boldsymbol{F}_s - \boldsymbol{p}) \\ \boldsymbol{F}'_k &= a\gamma(\boldsymbol{F}_s - \boldsymbol{p}) \end{aligned} \quad \text{where} \quad \begin{aligned} \gamma &= \frac{b}{|\boldsymbol{r}_{ij} + a\boldsymbol{r}_{jk}|} \\ \boldsymbol{p} &= \frac{\boldsymbol{r}_{is}\cdot\boldsymbol{F}_s}{\boldsymbol{r}_{is}\cdot\boldsymbol{r}_{is}}\boldsymbol{r}_{is} \end{aligned} \tag{4.152}$$

3fad. In the plane of three atoms, with a fixed angle and distance (Fig. 4.18 3fad):

$$\boldsymbol{r}_s = \boldsymbol{r}_i + d\cos\theta\frac{\boldsymbol{r}_{ij}}{|\boldsymbol{r}_{ij}|} + d\sin\theta\frac{\boldsymbol{r}_\perp}{|\boldsymbol{r}_\perp|} \quad \text{where} \quad \boldsymbol{r}_\perp = \boldsymbol{r}_{jk} - \frac{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{jk}}{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{ij}}\boldsymbol{r}_{ij} \tag{4.153}$$

In this case the virtual site is in the plane of the other three particles at a distance of $|d|$ from i at an angle of α with $\boldsymbol{r}_{ij}$. Atom k defines the plane and the direction of the angle. **Note** that in this case b and α must be specified, instead of a and b (see also sec. 5.2.2). The force on particles i, j and k due to the force on the virtual site can be computed as (with $\boldsymbol{r}_\perp$ as defined in eqn. 4.153):

$$\begin{aligned} \boldsymbol{F}'_i &= \boldsymbol{F}_s - \frac{d\cos\theta}{|\boldsymbol{r}_{ij}|}\boldsymbol{F}_1 + \frac{d\sin\theta}{|\boldsymbol{r}_\perp|}\left(\frac{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{jk}}{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{ij}}\boldsymbol{F}_2 + \boldsymbol{F}_3\right) \\ \boldsymbol{F}'_j &= \frac{d\cos\theta}{|\boldsymbol{r}_{ij}|}\boldsymbol{F}_1 - \frac{d\sin\theta}{|\boldsymbol{r}_\perp|}\left(\boldsymbol{F}_2 + \frac{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{jk}}{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{ij}}\boldsymbol{F}_2 + \boldsymbol{F}_3\right) \\ \boldsymbol{F}'_k &= \frac{d\sin\theta}{|\boldsymbol{r}_\perp|}\boldsymbol{F}_2 \end{aligned}$$

$$\text{where } \boldsymbol{F}_1 = \boldsymbol{F}_s - \frac{\boldsymbol{r}_{ij}\cdot\boldsymbol{F}_s}{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{ij}}\boldsymbol{r}_{ij}\,,\ \boldsymbol{F}_2 = \boldsymbol{F}_1 - \frac{\boldsymbol{r}_\perp\cdot\boldsymbol{F}_s}{\boldsymbol{r}_\perp\cdot\boldsymbol{r}_\perp}\boldsymbol{r}_\perp \text{ and } \boldsymbol{F}_3 = \frac{\boldsymbol{r}_{ij}\cdot\boldsymbol{F}_s}{\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{ij}}\boldsymbol{r}_\perp \tag{4.154}$$

3out. As a non-linear combination of three atoms, out of plane (Fig. 4.18 3out):

$$\boldsymbol{r}_s = \boldsymbol{r}_i + a\boldsymbol{r}_{ij} + b\boldsymbol{r}_{ik} + c(\boldsymbol{r}_{ij}\times\boldsymbol{r}_{ik}) \tag{4.155}$$

This enables the construction of virtual sites out of the plane of the other atoms. The force on particles i, j and k due to the force on the virtual site can be computed as:

$$\begin{aligned} \boldsymbol{F}'_j &= \begin{bmatrix} a & -c\,z_{ik} & c\,y_{ik} \\ c\,z_{ik} & a & -c\,x_{ik} \\ -c\,y_{ik} & c\,x_{ik} & a \end{bmatrix}\boldsymbol{F}_s \\ \boldsymbol{F}'_k &= \begin{bmatrix} b & c\,z_{ij} & -c\,y_{ij} \\ -c\,z_{ij} & b & c\,x_{ij} \\ c\,y_{ij} & -c\,x_{ij} & b \end{bmatrix}\boldsymbol{F}_s \\ \boldsymbol{F}'_i &= \boldsymbol{F}_s - \boldsymbol{F}'_j - \boldsymbol{F}'_k \end{aligned} \tag{4.156}$$

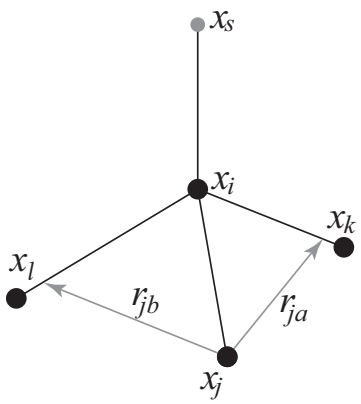

Figure 4.19: The new 4fdn virtual site construction, which is stable even when all constructing atoms are in the same plane.

4fdn. From four atoms, with a fixed distance, see separate Fig. 4.19. This construction is a bit complex, in particular since the previous type (4fd) could be unstable which forced us to introduce a more elaborate construction:

$$\begin{aligned}
\mathbf{r}_{ja} &= a\,\mathbf{r}_{ik} - \mathbf{r}_{ij} = a\,(\mathbf{x}_k - \mathbf{x}_i) - (\mathbf{x}_j - \mathbf{x}_i) \\
\mathbf{r}_{jb} &= b\,\mathbf{r}_{il} - \mathbf{r}_{ij} = b\,(\mathbf{x}_l - \mathbf{x}_i) - (\mathbf{x}_j - \mathbf{x}_i) \\
\mathbf{r}_m &= \mathbf{r}_{ja} \times \mathbf{r}_{jb} \\
\mathbf{x}_s &= \mathbf{x}_i + c\frac{\mathbf{r}_m}{|\mathbf{r}_m|}
\end{aligned} \tag{4.157}$$

In this case the virtual site is at a distance of $|c|$ from i, while a and b are parameters. **Note** that the vectors $\mathbf{r}_{ik}$ and $\mathbf{r}_{ij}$ are not normalized to save floating-point operations. The force on particles i, j, k and l due to the force on the virtual site are computed through chain rule derivatives of the construction expression. This is exact and conserves energy, but it does lead to relatively lengthy expressions that we do not include here (over 200 floating-point operations). The interested reader can look at the source code in `vsite.c`. Fortunately, this vsite type is normally only used for chiral centers such as C_α atoms in proteins.

The new 4fdn construct is identified with a 'type' value of 2 in the topology. The earlier 4fd type is still supported internally ('type' value 1), but it should not be used for new simulations. All current GROMACS tools will automatically generate type 4fdn instead.

N. A linear combination of N atoms with relative weights a_i. The weight for atom i is:

$$w_i = a_i \left(\sum_{j=1}^{N} a_j \right)^{-1} \tag{4.158}$$

There are three options for setting the weights:

COG center of geometry: equal weights

COM center of mass: a_i is the mass of atom i; when in free-energy simulations the mass of the atom is changed, only the mass of the A-state is used for the weight

COW center of weights: a_i is defined by the user

4.8 Long Range Electrostatics

4.8.1 Ewald summation

The total electrostatic energy of N particles and their periodic images is given by

$$V = \frac{f}{2} \sum_{n_x} \sum_{n_y} \sum_{n_z*} \sum_{i}^{N} \sum_{j}^{N} \frac{q_i q_j}{\mathbf{r}_{ij,\mathbf{n}}}. \tag{4.159}$$

$(n_x, n_y, n_z) = \mathbf{n}$ is the box index vector, and the star indicates that terms with $i = j$ should be omitted when $(n_x, n_y, n_z) = (0, 0, 0)$. The distance $\mathbf{r}_{ij,\mathbf{n}}$ is the real distance between the charges and not the minimum-image. This sum is conditionally convergent, but very slow.

Ewald summation was first introduced as a method to calculate long-range interactions of the periodic images in crystals [106]. The idea is to convert the single slowly-converging sum eqn. 4.159 into two quickly-converging terms and a constant term:

$$V = V_{\text{dir}} + V_{\text{rec}} + V_0 \tag{4.160}$$

$$V_{\text{dir}} = \frac{f}{2} \sum_{i,j}^{N} \sum_{n_x} \sum_{n_y} \sum_{n_z*} q_i q_j \frac{\text{erfc}(\beta r_{ij,\mathbf{n}})}{r_{ij,\mathbf{n}}} \tag{4.161}$$

$$V_{\text{rec}} = \frac{f}{2\pi V} \sum_{i,j}^{N} q_i q_j \sum_{m_x} \sum_{m_y} \sum_{m_z*} \frac{\exp\left(-(\pi \mathbf{m}/\beta)^2 + 2\pi i \mathbf{m} \cdot (\mathbf{r}_i - \mathbf{r}_j)\right)}{\mathbf{m}^2} \tag{4.162}$$

$$V_0 = -\frac{f\beta}{\sqrt{\pi}} \sum_{i}^{N} q_i^2, \tag{4.163}$$

where β is a parameter that determines the relative weight of the direct and reciprocal sums and $\mathbf{m} = (m_x, m_y, m_z)$. In this way we can use a short cut-off (of the order of 1 nm) in the direct space sum and a short cut-off in the reciprocal space sum (*e.g.* 10 wave vectors in each direction). Unfortunately, the computational cost of the reciprocal part of the sum increases as N^2 (or $N^{3/2}$ with a slightly better algorithm) and it is therefore not realistic for use in large systems.

Using Ewald

Don't use Ewald unless you are absolutely sure this is what you want - for almost all cases the PME method below will perform much better. If you still want to employ classical Ewald summation enter this in your `.mdp` file, if the side of your box is about 3 nm:

```
coulombtype      = Ewald
rvdw             = 0.9
```

```
rlist            = 0.9
rcoulomb         = 0.9
fourierspacing   = 0.6
ewald-rtol       = 1e-5
```

The ratio of the box dimensions and the `fourierspacing` parameter determines the highest magnitude of wave vectors m_x, m_y, m_z to use in each direction. With a 3-nm cubic box this example would use 11 wave vectors (from -5 to 5) in each direction. The `ewald-rtol` parameter is the relative strength of the electrostatic interaction at the cut-off. Decreasing this gives you a more accurate direct sum, but a less accurate reciprocal sum.

4.8.2 PME

Particle-mesh Ewald is a method proposed by Tom Darden [14] to improve the performance of the reciprocal sum. Instead of directly summing wave vectors, the charges are assigned to a grid using interpolation. The implementation in GROMACS uses cardinal B-spline interpolation [15], which is referred to as smooth PME (SPME). The grid is then Fourier transformed with a 3D FFT algorithm and the reciprocal energy term obtained by a single sum over the grid in k-space.

The potential at the grid points is calculated by inverse transformation, and by using the interpolation factors we get the forces on each atom.

The PME algorithm scales as $N \log(N)$, and is substantially faster than ordinary Ewald summation on medium to large systems. On very small systems it might still be better to use Ewald to avoid the overhead in setting up grids and transforms. For the parallelization of PME see the section on MPMD PME (3.17.5).

With the Verlet cut-off scheme, the PME direct space potential is shifted by a constant such that the potential is zero at the cut-off. This shift is small and since the net system charge is close to zero, the total shift is very small, unlike in the case of the Lennard-Jones potential where all shifts add up. We apply the shift anyhow, such that the potential is the exact integral of the force.

Using PME

As an example for using Particle-mesh Ewald summation in GROMACS, specify the following lines in your `.mdp` file:

```
coulombtype      = PME
rvdw             = 0.9
rlist            = 0.9
rcoulomb         = 0.9
fourierspacing   = 0.12
pme-order        = 4
ewald-rtol       = 1e-5
```

In this case the `fourierspacing` parameter determines the maximum spacing for the FFT grid (i.e. minimum number of grid points), and `pme-order` controls the interpolation order. Using

fourth-order (cubic) interpolation and this spacing should give electrostatic energies accurate to about $5 \cdot 10^{-3}$. Since the Lennard-Jones energies are not this accurate it might even be possible to increase this spacing slightly.

Pressure scaling works with PME, but be aware of the fact that anisotropic scaling can introduce artificial ordering in some systems.

4.8.3 P3M-AD

The Particle-Particle Particle-Mesh methods of Hockney & Eastwood can also be applied in GROMACS for the treatment of long range electrostatic interactions [107]. Although the P3M method was the first efficient long-range electrostatics method for molecular simulation, the smooth PME (SPME) method has largely replaced P3M as the method of choice in atomistic simulations. One performance disadvantage of the original P3M method was that it required 3 3D-FFT back transforms to obtain the forces on the particles. But this is not required for P3M and the forces can be derived through analytical differentiation of the potential, as done in PME. The resulting method is termed P3M-AD. The only remaining difference between P3M-AD and PME is the optimization of the lattice Green influence function for error minimization that P3M uses. However, in 2012 it has been shown that the SPME influence function can be modified to obtain P3M [108]. This means that the advantage of error minimization in P3M-AD can be used at the same computational cost and with the same code as PME, just by adding a few lines to modify the influence function. However, at optimal parameter setting the effect of error minimization in P3M-AD is less than 10%. P3M-AD does show large accuracy gains with interlaced (also known as staggered) grids, but that is not supported in GROMACS (yet).

P3M is used in GROMACS with exactly the same options as used with PME by selecting the electrostatics type:

```
coulombtype      = P3M-AD
```

4.8.4 Optimizing Fourier transforms and PME calculations

It is recommended to optimize the parameters for calculation of electrostatic interaction such as PME grid dimensions and cut-off radii. This is particularly relevant to do before launching long production runs.

`gmx mdrun` will automatically do a lot of PME optimization, and GROMACS also includes a special tool, `gmx tune_pme`, which automates the process of selecting the optimal number of PME-only ranks.

4.9 Long Range Van der Waals interactions

4.9.1 Dispersion correction

In this section, we derive long-range corrections due to the use of a cut-off for Lennard-Jones or Buckingham interactions. We assume that the cut-off is so long that the repulsion term can

safely be neglected, and therefore only the dispersion term is taken into account. Due to the nature of the dispersion interaction (we are truncating a potential proportional to $-r^{-6}$), energy and pressure corrections are both negative. While the energy correction is usually small, it may be important for free energy calculations where differences between two different Hamiltonians are considered. In contrast, the pressure correction is very large and can not be neglected under any circumstances where a correct pressure is required, especially for any NPT simulations. Although it is, in principle, possible to parameterize a force field such that the pressure is close to the desired experimental value without correction, such a method makes the parameterization dependent on the cut-off and is therefore undesirable.

Energy

The long-range contribution of the dispersion interaction to the virial can be derived analytically, if we assume a homogeneous system beyond the cut-off distance r_c. The dispersion energy between two particles is written as:

$$V(r_{ij}) \;=\; -C_6\, r_{ij}^{-6} \tag{4.164}$$

and the corresponding force is:

$$\boldsymbol{F}_{ij} \;=\; -6\, C_6\, r_{ij}^{-8} \boldsymbol{r}_{ij} \tag{4.165}$$

In a periodic system it is not easy to calculate the full potentials, so usually a cut-off is applied, which can be abrupt or smooth. We will call the potential and force with cut-off V_c and $\boldsymbol{F}_c$. The long-range contribution to the dispersion energy in a system with N particles and particle density $\rho = N/V$ is:

$$V_{lr} \;=\; \frac{1}{2} N \rho \int_0^\infty 4\pi r^2 g(r)\, (V(r) - V_c(r))\, \mathrm{d}r \tag{4.166}$$

We will integrate this for the shift function, which is the most general form of van der Waals interaction available in GROMACS. The shift function has a constant difference S from 0 to r_1 and is 0 beyond the cut-off distance r_c. We can integrate eqn. 4.166, assuming that the density in the sphere within r_1 is equal to the global density and the radial distribution function $g(r)$ is 1 beyond r_1:

$$\begin{aligned} V_{lr} &= \frac{1}{2} N \left(\rho \int_0^{r_1} 4\pi r^2 g(r)\, C_6\, S\, \mathrm{d}r + \rho \int_{r_1}^{r_c} 4\pi r^2\, (V(r) - V_c(r))\, \mathrm{d}r + \rho \int_{r_c}^{\infty} 4\pi r^2 V(r)\, \mathrm{d}r \right) \\ &= \frac{1}{2} N \left(\left(\frac{4}{3}\pi\rho r_1^3 - 1 \right) C_6\, S + \rho \int_{r_1}^{r_c} 4\pi r^2\, (V(r) - V_c(r))\, \mathrm{d}r - \frac{4}{3}\pi N \rho\, C_6\, r_c^{-3} \right) \end{aligned} \tag{4.167}$$

where the term -1 corrects for the self-interaction. For a plain cut-off we only need to assume that $g(r)$ is 1 beyond r_c and the correction reduces to [109]:

$$V_{lr} \;=\; -\frac{2}{3}\pi N \rho\, C_6\, r_c^{-3} \tag{4.168}$$

If we consider, for example, a box of pure water, simulated with a cut-off of 0.9 nm and a density of 1 g cm^{-3} this correction is -0.75 kJ mol^{-1} per molecule.

For a homogeneous mixture we need to define an *average dispersion constant*:

$$\langle C_6 \rangle = \frac{2}{N(N-1)} \sum_i^N \sum_{j>i}^N C_6(i,j) \tag{4.169}$$

In GROMACS, excluded pairs of atoms do not contribute to the average.

In the case of inhomogeneous simulation systems, *e.g.* a system with a lipid interface, the energy correction can be applied if $\langle C_6 \rangle$ for both components is comparable.

Virial and pressure

The scalar virial of the system due to the dispersion interaction between two particles i and j is given by:

$$\Xi \;=\; -\frac{1}{2}\boldsymbol{r}_{ij} \cdot \boldsymbol{F}_{ij} \;=\; 3\, C_6\, r_{ij}^{-6} \tag{4.170}$$

The pressure is given by:

$$P \;=\; \frac{2}{3\,V}\,(E_{kin} - \Xi) \tag{4.171}$$

The long-range correction to the virial is given by:

$$\Xi_{lr} \;=\; \frac{1}{2} N \rho \int_0^\infty 4\pi r^2 g(r)(\Xi - \Xi_c)\,\mathrm{d}r \tag{4.172}$$

We can again integrate the long-range contribution to the virial assuming $g(r)$ is 1 beyond r_1:

$$\begin{aligned} \Xi_{lr} \;&=\; \frac{1}{2} N \rho \left(\int_{r_1}^{r_c} 4\pi r^2 (\Xi - \Xi_c)\,\mathrm{d}r + \int_{r_c}^{\infty} 4\pi r^2 3\, C_6\, r_{ij}^{-6}\,\mathrm{d}r \right) \\ &=\; \frac{1}{2} N \rho \left(\int_{r_1}^{r_c} 4\pi r^2 (\Xi - \Xi_c)\,\mathrm{d}r + 4\pi C_6\, r_c^{-3} \right) \end{aligned} \tag{4.173}$$

For a plain cut-off the correction to the pressure is [109]:

$$P_{lr} \;=\; -\frac{4}{3}\pi C_6\, \rho^2 r_c^{-3} \tag{4.174}$$

Using the same example of a water box, the correction to the virial is 0.75 kJ mol^{-1} per molecule, the corresponding correction to the pressure for SPC water is approximately -280 bar.

For homogeneous mixtures, we can again use the average dispersion constant $\langle C_6 \rangle$ (eqn. 4.169):

$$P_{lr} \;=\; -\frac{4}{3}\pi \,\langle C_6 \rangle\, \rho^2 r_c^{-3} \tag{4.175}$$

For inhomogeneous systems, eqn. 4.175 can be applied under the same restriction as holds for the energy (see sec. 4.9.1).

4.9.2 Lennard-Jones PME

In order to treat systems, using Lennard-Jones potentials, that are non-homogeneous outside of the cut-off distance, we can instead use the Particle-mesh Ewald method as discussed for electrostatics above. In this case the modified Ewald equations become

$$V \;=\; V_{\mathrm{dir}} + V_{\mathrm{rec}} + V_0 \tag{4.176}$$

$$V_{\text{dir}} = -\frac{1}{2}\sum_{i,j}^{N}\sum_{n_x}\sum_{n_y}\sum_{n_z*}\frac{C_6^{ij}\,g(\beta r_{ij,\mathbf{n}})}{{r_{ij,\mathbf{n}}}^6} \tag{4.177}$$

$$V_{\text{rec}} = \frac{\pi^{\frac{3}{2}}\beta^3}{2V}\sum_{m_x}\sum_{m_y}\sum_{m_z*} f(\pi|\mathbf{m}|/\beta)\times\sum_{i,j}^{N} C_6^{ij}\exp\left[-2\pi i\mathbf{m}\cdot(\mathbf{r_i}-\mathbf{r_j})\right] \tag{4.178}$$

$$V_0 = -\frac{\beta^6}{12}\sum_{i}^{N} C_6^{ii} \tag{4.179}$$

where $\mathbf{m} = (m_x, m_y, m_z)$, β is the parameter determining the weight between direct and reciprocal space, and C_6^{ij} is the combined dispersion parameter for particle i and j. The star indicates that terms with $i = j$ should be omitted when $((n_x, n_y, n_z) = (0, 0, 0))$, and $\mathbf{r}_{ij,\mathbf{n}}$ is the real distance between the particles. Following the derivation by Essmann [15], the functions f and g introduced above are defined as

$$f(x) = 1/3\left[(1-2x^2)\exp(-x^2) + 2x^3\sqrt{\pi}\,\text{erfc}(x)\right] \tag{4.180}$$

$$g(x) = \exp(-x^2)(1 + x^2 + \frac{x^4}{2}). \tag{4.181}$$

The above methodology works fine as long as the dispersion parameters can be combined geometrically (eqn. 4.6) in the same way as the charges for electrostatics

$$C_{6,\text{geom}}^{ij} = \left(C_6^{ii}\,C_6^{jj}\right)^{1/2} \tag{4.182}$$

For Lorentz-Berthelot combination rules (eqn. 4.7), the reciprocal part of this sum has to be calculated seven times due to the splitting of the dispersion parameter according to

$$C_{6,\text{L-B}}^{ij} = (\sigma_i + \sigma_j)^6 = \sum_{n=0}^{6} P_n \sigma_i^n \sigma_j^{(6-n)}, \tag{4.183}$$

for P_n the Pascal triangle coefficients. This introduces a non-negligible cost to the reciprocal part, requiring seven separate FFTs, and therefore this has been the limiting factor in previous attempts to implement LJ-PME. A solution to this problem is to use geometrical combination rules in order to calculate an approximate interaction parameter for the reciprocal part of the potential, yielding a total interaction of

$$V(r < r_c) = \underbrace{C_6^{\text{dir}} g(\beta r) r^{-6}}_{\text{Direct space}} + \underbrace{C_{6,\text{geom}}^{\text{recip}}[1 - g(\beta r)] r^{-6}}_{\text{Reciprocal space}}$$

$$= C_{6,\text{geom}}^{\text{recip}} r^{-6} + \left(C_6^{\text{dir}} - C_{6,\text{geom}}^{\text{recip}}\right) g(\beta r) r^{-6} \tag{4.184}$$

$$V(r > r_c) = \underbrace{C_{6,\text{geom}}^{\text{recip}}[1 - g(\beta r)] r^{-6}}_{\text{Reciprocal space}}. \tag{4.185}$$

This will preserve a well-defined Hamiltonian and significantly increase the performance of the simulations. The approximation does introduce some errors, but since the difference is located in the interactions calculated in reciprocal space, the effect will be very small compared to the total

interaction energy. In a simulation of a lipid bilayer, using a cut-off of 1.0 nm, the relative error in total dispersion energy was below 0.5%. A more thorough discussion of this can be found in [110].

In GROMACS we now perform the proper calculation of this interaction by subtracting, from the direct-space interactions, the contribution made by the approximate potential that is used in the reciprocal part

$$V_{\text{dir}} = C_6^{\text{dir}} r^{-6} - C_6^{\text{recip}}[1 - g(\beta r)]r^{-6}. \tag{4.186}$$

This potential will reduce to the expression in eqn. 4.177 when $C_6^{\text{dir}} = C_6^{\text{recip}}$, and the total interaction is given by

$$\begin{aligned} V(r < r_c) &= \underbrace{C_6^{\text{dir}} r^{-6} - C_6^{\text{recip}}[1 - g(\beta r)]r^{-6}}_{\text{Direct space}} + \underbrace{C_6^{\text{recip}}[1 - g(\beta r)]r^{-6}}_{\text{Reciprocal space}} \\ &= C_6^{\text{dir}} r^{-6} \qquad (4.187) \\ V(r > r_c) &= C_6^{\text{recip}}[1 - g(\beta r)]r^{-6}. \qquad (4.188) \end{aligned}$$

For the case when $C_6^{\text{dir}} \neq C_6^{\text{recip}}$ this will retain an unmodified LJ force up to the cut-off, and the error is an order of magnitude smaller than in simulations where the direct-space interactions do not account for the approximation used in reciprocal space. When using a VdW interaction modifier of potential-shift, the constant

$$\left(-C_6^{\text{dir}} + C_6^{\text{recip}}[1 - g(\beta r_c)]\right) r_c^{-6} \tag{4.189}$$

is added to eqn. 4.187 in order to ensure that the potential is continuous at the cutoff. Note that, in the same way as eqn. 4.186, this degenerates into the expected $-C_6 g(\beta r_c) r_c^{-6}$ when $C_6^{\text{dir}} = C_6^{\text{recip}}$. In addition to this, a long-range dispersion correction can be applied to correct for the approximation using a combination rule in reciprocal space. This correction assumes, as for the cut-off LJ potential, a uniform particle distribution. But since the error of the combination rule approximation is very small this long-range correction is not necessary in most cases. Also note that this homogenous correction does not correct the surface tension, which is an inhomogeneous property.

Using LJ-PME

As an example for using Particle-mesh Ewald summation for Lennard-Jones interactions in GROMACS, specify the following lines in your `.mdp` file:

```
vdwtype           = PME
rvdw              = 0.9
vdw-modifier      = Potential-Shift
rlist             = 0.9
rcoulomb          = 0.9
fourierspacing    = 0.12
pme-order         = 4
ewald-rtol-lj     = 0.001
lj-pme-comb-rule = geometric
```

The same Fourier grid and interpolation order are used if both LJ-PME and electrostatic PME are active, so the settings for `fourierspacing` and `pme-order` are common to both. `ewald-rtol-lj` controls the splitting between direct and reciprocal space in the same way as `ewald-rtol`. In addition to this, the combination rule to be used in reciprocal space is determined by `lj-pme-comb-rule`. If the current force field uses Lorentz-Berthelot combination rules, it is possible to set `lj-pme-comb-rule = geometric` in order to gain a significant increase in performance for a small loss in accuracy. The details of this approximation can be found in the section above.

Note that the use of a complete long-range dispersion correction means that as with Coulomb PME, `rvdw` is now a free parameter in the method, rather than being necessarily restricted by the force-field parameterization scheme. Thus it is now possible to optimize the cutoff, spacing, order and tolerance terms for accuracy and best performance.

Naturally, the use of LJ-PME rather than LJ cut-off adds computation and communication done for the reciprocal-space part, so for best performance in balancing the load of parallel simulations using PME-only ranks, more such ranks should be used. It may be possible to improve upon the automatic load-balancing used by `mdrun`.

4.10 Force field

A force field is built up from two distinct components:

- The set of equations (called the *potential functions*) used to generate the potential energies and their derivatives, the forces. These are described in detail in the previous chapter.
- The parameters used in this set of equations. These are not given in this manual, but in the data files corresponding to your GROMACS distribution.

Within one set of equations various sets of parameters can be used. Care must be taken that the combination of equations and parameters form a consistent set. It is in general dangerous to make *ad hoc* changes in a subset of parameters, because the various contributions to the total force are usually interdependent. This means in principle that every change should be documented, verified by comparison to experimental data and published in a peer-reviewed journal before it can be used.

GROMACS 2016 includes several force fields, and additional ones are available on the website. If you do not know which one to select we recommend GROMOS-96 for united-atom setups and OPLS-AA/L for all-atom parameters. That said, we describe the available options in some detail.

All-hydrogen force field

The GROMOS-87-based all-hydrogen force field is almost identical to the normal GROMOS-87 force field, since the extra hydrogens have no Lennard-Jones interaction and zero charge. The only differences are in the bond angle and improper dihedral angle terms. This force field is only useful when you need the exact hydrogen positions, for instance for distance restraints derived from NMR measurements. When citing this force field please read the previous paragraph.

4.10.1 GROMOS-96

GROMACS supports the GROMOS-96 force fields [81]. All parameters for the 43A1, 43A2 (development, improved alkane dihedrals), 45A3, 53A5, and 53A6 parameter sets are included. All standard building blocks are included and topologies can be built automatically by `pdb2gmx`.

The GROMOS-96 force field is a further development of the GROMOS-87 force field. It has improvements over the GROMOS-87 force field for proteins and small molecules. **Note** that the sugar parameters present in 53A6 do correspond to those published in 2004[111], which are different from those present in 45A4, which is not included in GROMACS at this time. The 45A4 parameter set corresponds to a later revision of these parameters. The GROMOS-96 force field is not, however, recommended for use with long alkanes and lipids. The GROMOS-96 force field differs from the GROMOS-87 force field in a few respects:

- the force field parameters
- the parameters for the bonded interactions are not linked to atom types
- a fourth power bond stretching potential (4.2.1)
- an angle potential based on the cosine of the angle (4.2.6)

There are two differences in implementation between GROMACS and GROMOS-96 which can lead to slightly different results when simulating the same system with both packages:

- in GROMOS-96 neighbor searching for solvents is performed on the first atom of the solvent molecule. This is not implemented in GROMACS, but the difference with searching by centers of charge groups is very small
- the virial in GROMOS-96 is molecule-based. This is not implemented in GROMACS, which uses atomic virials

The GROMOS-96 force field was parameterized with a Lennard-Jones cut-off of 1.4 nm, so be sure to use a Lennard-Jones cut-off (`rvdw`) of at least 1.4. A larger cut-off is possible because the Lennard-Jones potential and forces are almost zero beyond 1.4 nm.

GROMOS-96 files

GROMACS can read and write GROMOS-96 coordinate and trajectory files. These files should have the extension `.g96`. Such a file can be a GROMOS-96 initial/final configuration file, a coordinate trajectory file, or a combination of both. The file is fixed format; all floats are written as 15.9, and as such, files can get huge. GROMACS supports the following data blocks in the given order:

- Header block:

```
TITLE (mandatory)
```

- Frame blocks:

```
TIMESTEP (optional)
POSITION/POSITIONRED (mandatory)
VELOCITY/VELOCITYRED (optional)
BOX (optional)
```

See the GROMOS-96 manual [81] for a complete description of the blocks. **Note** that all GROMACS programs can read compressed (.Z) or gzipped (.gz) files.

4.10.2 OPLS/AA

4.10.3 AMBER

GROMACS provides native support for the following AMBER force fields:

- AMBER94 [112]
- AMBER96 [113]
- AMBER99 [114]
- AMBER99SB [115]
- AMBER99SB-ILDN [116]
- AMBER03 [117]
- AMBERGS [118]

4.10.4 CHARMM

GROMACS supports the CHARMM force field for proteins [119, 120], lipids [121] and nucleic acids [122, 123]. The protein parameters (and to some extent the lipid and nucleic acid parameters) were thoroughly tested – both by comparing potential energies between the port and the standard parameter set in the CHARMM molecular simulation package, as well by how the protein force field behaves together with GROMACS-specific techniques such as virtual sites (enabling long time steps) and a fast implicit solvent recently implemented [73] – and the details and results are presented in the paper by Bjelkmar et al. [124]. The nucleic acid parameters, as well as the ones for HEME, were converted and tested by Michel Cuendet.

When selecting the CHARMM force field in `pdb2gmx` the default option is to use CMAP (for torsional correction map). To exclude CMAP, use `-nocmap`. The basic form of the CMAP term implemented in GROMACS is a function of the ϕ and ψ backbone torsion angles. This term is defined in the `.rtp` file by a `[ cmap ]` statement at the end of each residue supporting CMAP. The following five atom names define the two torsional angles. Atoms 1-4 define ϕ, and atoms 2-5 define ψ. The corresponding atom types are then matched to the correct CMAP type in the `cmap.itp` file that contains the correction maps.

A port of the CHARMM36 force field for use with GROMACS is also available at http://mackerell.umaryland.edu/charmm_ff.shtml#gromacs.

For branched polymers or other topologies not supported by pdb2gmx, it is possible to use TopoTools [125] to generate a GROMACS top file.

4.10.5 Coarse-grained force fields

Coarse-graining is a systematic way of reducing the number of degrees of freedom representing a system of interest. To achieve this, typically whole groups of atoms are represented by single beads and the coarse-grained force fields describes their effective interactions. Depending on the choice of parameterization, the functional form of such an interaction can be complicated and often tabulated potentials are used.

Coarse-grained models are designed to reproduce certain properties of a reference system. This can be either a full atomistic model or even experimental data. Depending on the properties to reproduce there are different methods to derive such force fields. An incomplete list of methods is given below:

- Conserving free energies
  - Simplex method
  - MARTINI force field (see next section)
- Conserving distributions (like the radial distribution function), so-called structure-based coarse-graining
  - (iterative) Boltzmann inversion
  - Inverse Monte Carlo
- Conversing forces
  - Force matching

Note that coarse-grained potentials are state dependent (e.g. temperature, density,...) and should be re-parametrized depending on the system of interest and the simulation conditions. This can for example be done using the Versatile Object-oriented Toolkit for Coarse-Graining Applications (VOTCA) [126]. The package was designed to assists in systematic coarse-graining, provides implementations for most of the algorithms mentioned above and has a well tested interface to GROMACS. It is available as open source and further information can be found at www.votca.org.

4.10.6 MARTINI

The MARTINI force field is a coarse-grain parameter set that allows for the construction of many systems, including proteins and membranes.

4.10.7 PLUM

The PLUM force field [127] is an example of a solvent-free protein-membrane model for which the membrane was derived from structure-based coarse-graining [128]. A GROMACS implementation can be found at code.google.com/p/plumx.