Chapter 6

Special Topics

6.1   Free energy implementation

For free energy calculations, there are two things that must be specified; the end states, and the pathway connecting the end states. The end states can be specified in two ways. The most straightforward is through the specification of end states in the topology file. Most potential forms support both an A state and a B state. Whenever both states are specified, then the A state corresponds to the initial free energy state, and the B state corresponds to the final state.

In some cases, the end state can also be defined in some cases without altering the topology, solely through the `.mdp` file, through the use of the `couple-moltype`,`couple-lambda0`, `couple-lambda1`, and `couple-intramol` mdp keywords. Any molecule type selected in `couple-moltype` will automatically have a B state implicitly constructed (and the A state redefined) according to the `couple-lambda` keywords. `couple-lambda0` and `couple-lambda1` define the non-bonded parameters that are present in the A state (`couple-lambda0`) and the B state (`couple-lambda1`). The choices are 'q','vdw', and 'vdw-q'; these indicate the Coulombic, van der Waals, or both parameters that are turned on in the respective state.

Once the end states are defined, then the path between the end states has to be defined. This path is defined solely in the .mdp file. Starting in 4.6, λ is a vector of components, with Coulombic, van der Waals, bonded, restraint, and mass components all able to be adjusted independently. This makes it possible to turn off the Coulombic term linearly, and then the van der Waals using soft core, all in the same simulation. This is especially useful for replica exchange or expanded ensemble simulations, where it is important to sample all the way from interacting to non-interacting states in the same simulation to improve sampling.

`fep-lambdas` is the default array of λ values ranging from 0 to 1. All of the other lambda arrays use the values in this array if they are not specified. The previous behavior, where the pathway is controlled by a single λ variable, can be preserved by using only `fep-lambdas` to define the pathway.

For example, if you wanted to first to change the Coulombic terms, then the van der Waals terms, changing bonded at the same time rate as the van der Waals, but changing the restraints throughout

the first two-thirds of the simulation, then you could use this λ vector:

```
coul-lambdas            = 0.0 0.2 0.5 1.0 1.0 1.0 1.0 1.0 1.0 1.0
vdw-lambdas             = 0.0 0.0 0.0 0.0 0.4 0.5 0.6 0.7 0.8 1.0
bonded-lambdas          = 0.0 0.0 0.0 0.0 0.4 0.5 0.6 0.7 0.8 1.0
restraint-lambdas       = 0.0 0.0 0.1 0.2 0.3 0.5 0.7 1.0 1.0 1.0
```

This is also equivalent to:

```
fep-lambdas             = 0.0 0.0 0.0 0.0 0.4 0.5 0.6 0.7 0.8 1.0
coul-lambdas            = 0.0 0.2 0.5 1.0 1.0 1.0 1.0 1.0 1.0 1.0
restraint-lambdas       = 0.0 0.0 0.1 0.2 0.3 0.5 0.7 1.0 1.0 1.0
```

The `fep-lambda array`, in this case, is being used as the default to fill in the bonded and van der Waals λ arrays. Usually, it's best to fill in all arrays explicitly, just to make sure things are properly assigned.

If you want to turn on only restraints going from A to B, then it would be:

```
restraint-lambdas       = 0.0 0.1 0.2 0.4 0.6 1.0
```

and all of the other components of the λ vector would be left in the A state.

To compute free energies with a vector λ using thermodynamic integration, then the TI equation becomes vector equation:

$$\Delta F = \int \langle \nabla H \rangle \cdot d\vec{\lambda} \tag{6.1}$$

or for finite differences:

$$\Delta F \approx \int \sum \langle \nabla H \rangle \cdot \Delta \lambda \tag{6.2}$$

The external `pymbar` script downloaded from https://SimTK.org/home/pymbar can compute this integral automatically from the GROMACS dhdl.xvg output.

6.2   Potential of mean force

A potential of mean force (PMF) is a potential that is obtained by integrating the mean force from an ensemble of configurations. In GROMACS, there are several different methods to calculate the mean force. Each method has its limitations, which are listed below.

- **pull code:** between the centers of mass of molecules or groups of molecules.

- **free-energy code with harmonic bonds or constraints:** between single atoms.

- **free-energy code with position restraints:** changing the conformation of a relatively immobile group of atoms.

- **pull code in limited cases:** between groups of atoms that are part of a larger molecule for which the bonds are constrained with SHAKE or LINCS. If the pull group if relatively large, the pull code can be used.

The pull and free-energy code a described in more detail in the following two sections.

Entropic effects

When a distance between two atoms or the centers of mass of two groups is constrained or re-strained, there will be a purely entropic contribution to the PMF due to the rotation of the two groups [131]. For a system of two non-interacting masses the potential of mean force is:

$$V_{pmf}(r) = -(n_c - 1)k_B T \log(r) \tag{6.3}$$

where n_c is the number of dimensions in which the constraint works (i.e. $n_c = 3$ for a normal con-straint and $n_c = 1$ when only the z-direction is constrained). Whether one needs to correct for this contribution depends on what the PMF should represent. When one wants to pull a substrate into a protein, this entropic term indeed contributes to the work to get the substrate into the protein. But when calculating a PMF between two solutes in a solvent, for the purpose of simulating without solvent, the entropic contribution should be removed. **Note** that this term can be significant; when at 300K the distance is halved, the contribution is 3.5 kJ mol^{-1}.

6.3 Non-equilibrium pulling

When the distance between two groups is changed continuously, work is applied to the system, which means that the system is no longer in equilibrium. Although in the limit of very slow pulling the system is again in equilibrium, for many systems this limit is not reachable within reasonable computational time. However, one can use the Jarzynski relation [132] to obtain the equilibrium free-energy difference ΔG between two distances from many non-equilibrium simulations:

$$\Delta G_{AB} = -k_B T \log \left\langle e^{-\beta W_{AB}} \right\rangle_A \tag{6.4}$$

where W_{AB} is the work performed to force the system along one path from state A to B, the angular bracket denotes averaging over a canonical ensemble of the initial state A and $\beta = 1/k_B T$.

6.4 The pull code

The pull code applies forces or constraints between the centers of mass of one or more pairs of groups of atoms. Each pull reaction coordinate is called a "coordinate" and it operates on two pull groups. A pull group can be part of one or more pull coordinates. Furthermore, a coordinate can also operate on a single group and an absolute reference position in space. The distance between a pair of groups can be determined in 1, 2 or 3 dimensions, or can be along a user-defined vector. The reference distance can be constant or can change linearly with time. Normally all atoms are weighted by their mass, but an additional weighting factor can also be used.
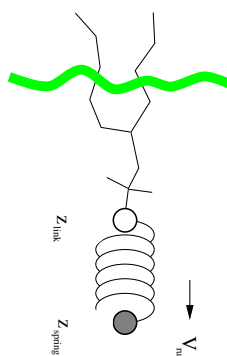
Figure 6.1: Schematic picture of pulling a lipid out of a lipid bilayer with umbrella pulling. V_{rup} is the velocity at which the spring is retracted, Z_{link} is the atom to which the spring is attached and Z_{spring} is the location of the spring.

Three different types of calculation are supported, and in all cases the reference distance can be constant or linearly changing with time.

1. **Umbrella pulling** A harmonic potential is applied between the centers of mass of two groups. Thus, the force is proportional to the displacement.

2. **Constraint pulling** The distance between the centers of mass of two groups is constrained. The constraint force can be written to a file. This method uses the SHAKE algorithm but only needs 1 iteration to be exact if only two groups are constrained.

3. **Constant force pulling** A constant force is applied between the centers of mass of two groups. Thus, the potential is linear. In this case there is no reference distance of pull rate.

Definition of the center of mass

In GROMACS, there are three ways to define the center of mass of a group. The standard way is a "plain" center of mass, possibly with additional weighting factors. With periodic boundary conditions it is no longer possible to uniquely define the center of mass of a group of atoms. Therefore, a reference atom is used. For determining the center of mass, for all other atoms in the group, the closest periodic image to the reference atom is used. This uniquely defines the center of mass. By default, the middle (determined by the order in the topology) atom is used as a reference atom, but the user can also select any other atom if it would be closer to center of the group.

For a layered system, for instance a lipid bilayer, it may be of interest to calculate the PMF of a lipid as function of its distance from the whole bilayer. The whole bilayer can be taken as reference group in that case, but it might also be of interest to define the reaction coordinate for the PMF more locally. The `.mdp` option `pull_geometry = cylinder` does not use all the atoms of the reference group, but instead dynamically only those within a cylinder with radius `r_1` around the pull vector going through the pull group. This only works for distances defined in
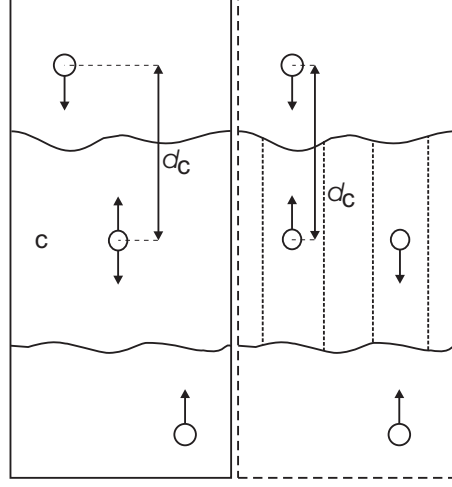
Figure 6.2: Comparison of a plain center of mass reference group versus a cylinder reference group applied to interface systems. C is the reference group. The circles represent the center of mass of two groups plus the reference group, d_c is the reference distance.

one dimension, and the cylinder is oriented with its long axis along this one dimension. A second cylinder can be defined with `r_0`, with a linear switch function that weighs the contribution of atoms between `r_0` and `r_1` with distance. This smooths the effects of atoms moving in and out of the cylinder (which causes jumps in the pull forces).

For a group of molecules in a periodic system, a plain reference group might not be well-defined. An example is a water slab that is connected periodically in x and y, but has two liquid-vapor interfaces along z. In such a setup, water molecules can evaporate from the liquid and they will move through the vapor, through the periodic boundary, to the other interface. Such a system is inherently periodic and there is no proper way of defining a "plain" center of mass along z. A proper solution is to using a cosine shaped weighting profile for all atoms in the reference group. The profile is a cosine with a single period in the unit cell. Its phase is optimized to give the maximum sum of weights, including mass weighting. This provides a unique and continuous reference position that is nearly identical to the plain center of mass position in case all atoms are all within a half of the unit-cell length. See ref [133] for details.

When relative weights w_i are used during the calculations, either by supplying weights in the input or due to cylinder geometry or due to cosine weighting, the weights need to be scaled to conserve momentum:

$$w'_i = w_i \sum_{j=1}^{N} w_j \, m_j \left/ \sum_{j=1}^{N} w_j^2 \, m_j \right. \tag{6.5}$$

where m_j is the mass of atom j of the group. The mass of the group, required for calculating the constraint force, is:

$$M = \sum_{i=1}^{N} w'_i \, m_i \tag{6.6}$$

The definition of the weighted center of mass is:

$$\boldsymbol{r}_{com} = \sum_{i=1}^{N} w_i' \, m_i \, \boldsymbol{r}_i \bigg/ M \tag{6.7}$$

From the centers of mass the AFM, constraint, or umbrella force $\boldsymbol{F}_{com}$ on each group can be calculated. The force on the center of mass of a group is redistributed to the atoms as follows:

$$\boldsymbol{F}_i = \frac{w_i' \, m_i}{M} \, \boldsymbol{F}_{com} \tag{6.8}$$

Limitations

There is one theoretical limitation: strictly speaking, constraint forces can only be calculated between groups that are not connected by constraints to the rest of the system. If a group contains part of a molecule of which the bond lengths are constrained, the pull constraint and LINCS or SHAKE bond constraint algorithms should be iterated simultaneously. This is not done in GROMACS. This means that for simulations with `constraints = all-bonds` in the `.mdp` file pulling is, strictly speaking, limited to whole molecules or groups of molecules. In some cases this limitation can be avoided by using the free energy code, see sec. 6.7. In practice, the errors caused by not iterating the two constraint algorithms can be negligible when the pull group consists of a large amount of atoms and/or the pull force is small. In such cases, the constraint correction displacement of the pull group is small compared to the bond lengths.

6.5   Enforced Rotation

This module can be used to enforce the rotation of a group of atoms, as *e.g.* a protein subunit. There are a variety of rotation potentials, among them complex ones that allow flexible adaptations of both the rotated subunit as well as the local rotation axis during the simulation. An example application can be found in ref. [134].

6.5.1   Fixed Axis Rotation

Stationary Axis with an Isotropic Potential

In the fixed axis approach (see Fig. 6.3B), torque on a group of N atoms with positions $\boldsymbol{x}_i$ (denoted "rotation group") is applied by rotating a reference set of atomic positions – usually their initial positions $\boldsymbol{y}_i^0$ – at a constant angular velocity ω around an axis defined by a direction vector $\hat{\boldsymbol{v}}$ and a pivot point $\boldsymbol{u}$. To that aim, each atom with position $\boldsymbol{x}_i$ is attracted by a "virtual spring" potential to its moving reference position $\boldsymbol{y}_i = \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u})$, where $\boldsymbol{\Omega}(t)$ is a matrix that describes the rotation around the axis. In the simplest case, the "springs" are described by a harmonic potential,

$$V^{\text{iso}} = \frac{k}{2} \sum_{i=1}^{N} w_i \left[\boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u}) - (\boldsymbol{x}_i - \boldsymbol{u}) \right]^2, \tag{6.9}$$
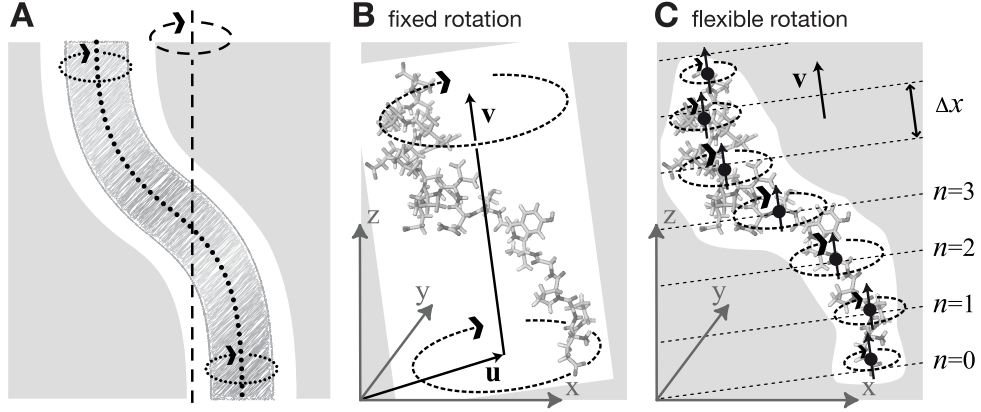
Figure 6.3: Comparison of fixed and flexible axis rotation. **A**: Rotating the sketched shape inside the white tubular cavity can create artifacts when a fixed rotation axis (dashed) is used. More realistically, the shape would revolve like a flexible pipe-cleaner (dotted) inside the bearing (gray). **B**: Fixed rotation around an axis v with a pivot point specified by the vector u. **C**: Subdividing the rotating fragment into slabs with separate rotation axes ($\uparrow$) and pivot points ($\bullet$) for each slab allows for flexibility. The distance between two slabs with indices n and $n + 1$ is Δx.

with optional mass-weighted prefactors $w_i = N m_i/M$ with total mass $M = \sum_{i=1}^{N} m_i$. The rotation matrix $\mathbf{\Omega}(t)$ is

$$\mathbf{\Omega}(t) = \begin{pmatrix} \cos\omega t + v_x^2 \xi & v_x v_y \xi - v_z \sin\omega t & v_x v_z \xi + v_y \sin\omega t \\ v_x v_y \xi + v_z \sin\omega t & \cos\omega t + v_y^2 \xi & v_y v_z \xi - v_x \sin\omega t \\ v_x v_z \xi - v_y \sin\omega t & v_y v_z \xi + v_x \sin\omega t & \cos\omega t + v_z^2 \xi \end{pmatrix},$$

where v_x, v_y, and v_z are the components of the normalized rotation vector $\hat{v}$, and $\xi := 1 - \cos(\omega t)$. As illustrated in Fig. 6.4A for a single atom j, the rotation matrix $\mathbf{\Omega}(t)$ operates on the initial reference positions $y_j^0 = x_j(t_0)$ of atom j at $t = t_0$. At a later time t, the reference position has rotated away from its initial place (along the blue dashed line), resulting in the force

$$\boldsymbol{F}_j^{\text{iso}} = -\nabla_j V^{\text{iso}} = k\, w_j \left[\mathbf{\Omega}(t)(y_j^0 - u) - (x_j - u) \right], \tag{6.10}$$

which is directed towards the reference position.

Pivot-Free Isotropic Potential

Instead of a fixed pivot vector u this potential uses the center of mass x_c of the rotation group as pivot for the rotation axis,

$$x_c = \frac{1}{M} \sum_{i=1}^{N} m_i x_i \qquad \text{and} \qquad y_c^0 = \frac{1}{M} \sum_{i=1}^{N} m_i y_i^0, \tag{6.11}$$

which yields the "pivot-free" isotropic potential

$$V^{\text{iso-pf}} = \frac{k}{2} \sum_{i=1}^{N} w_i \left[\mathbf{\Omega}(t)(y_i^0 - y_c^0) - (x_i - x_c) \right]^2, \tag{6.12}$$

A V^{iso}

B $V^{\mathrm{rm}}, V^{\mathrm{flex}}$

C $V^{\mathrm{rm2}}, V^{\mathrm{flex2}}$ ($\epsilon' = 0 \ \mathrm{nm}^2$)

D $V^{\mathrm{rm2}}, V^{\mathrm{flex2}}$ ($\epsilon' = 0.01 \ \mathrm{nm}^2$)
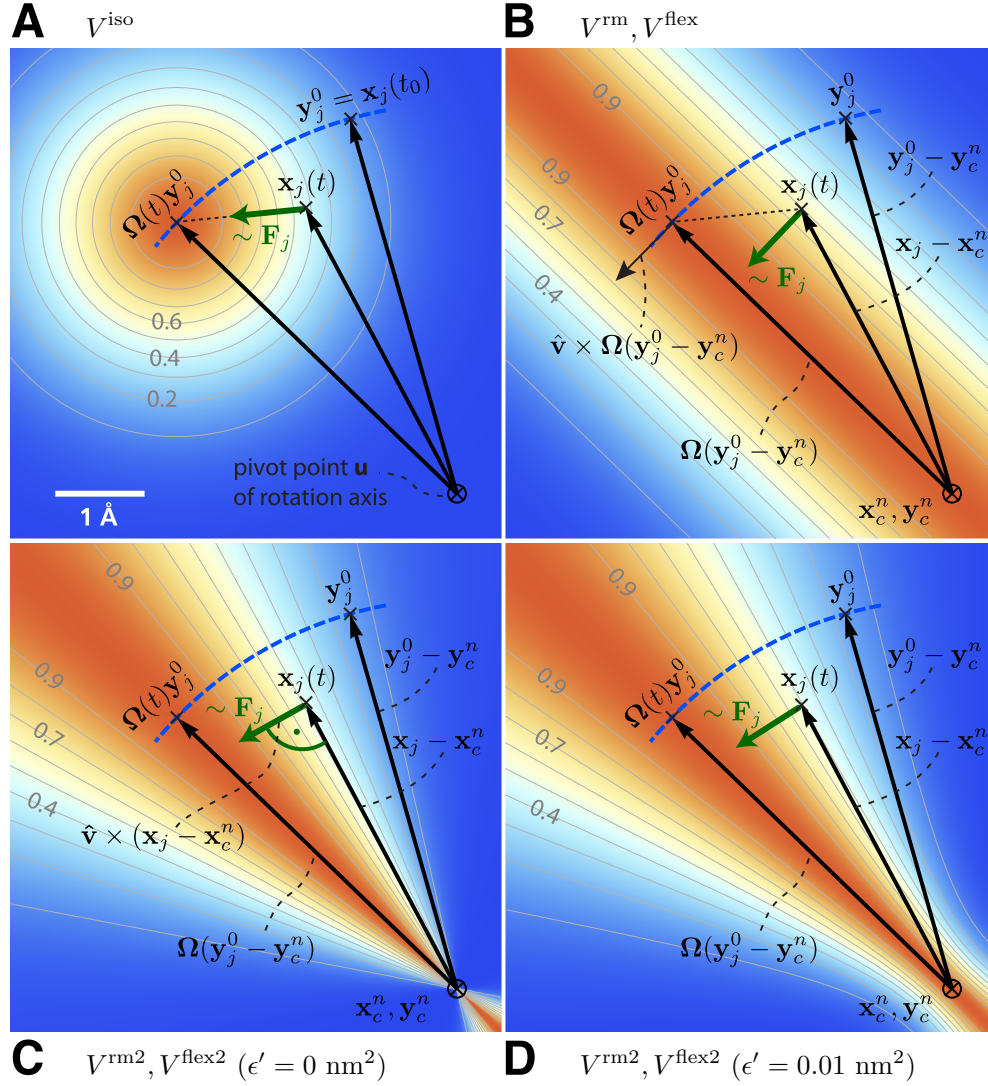
Figure 6.4: Selection of different rotation potentials and definition of notation. All four potentials V (color coded) are shown for a single atom at position $\boldsymbol{x}_j(t)$. **A**: Isotropic potential V^{iso}, **B**: radial motion potential V^{rm} and flexible potential V^{flex}, **C–D**: radial motion 2 potential V^{rm2} and flexible 2 potential V^{flex2} for $\epsilon' = 0 \ \mathrm{nm}^2$ (**C**) and $\epsilon' = 0.01 \ \mathrm{nm}^2$ (**D**). The rotation axis is perpendicular to the plane and marked by $\otimes$. The light gray contours indicate Boltzmann factors $e^{-V/(k_B T)}$ in the $\boldsymbol{x}_j$-plane for $T = 300 \ \mathrm{K}$ and $k = 200 \ \mathrm{kJ/(mol \cdot nm^2)}$. The green arrow shows the direction of the force $\boldsymbol{F}_j$ acting on atom j; the blue dashed line indicates the motion of the reference position.

with forces

$$\mathbf{F}_j^{\text{iso-pf}} = k\, w_j \left[\mathbf{\Omega}(t)(\mathbf{y}_j^0 - \mathbf{y}_c^0) - (\mathbf{x}_j - \mathbf{x}_c) \right]. \tag{6.13}$$

Without mass-weighting, the pivot $\mathbf{x}_c$ is the geometrical center of the group.

Parallel Motion Potential Variant

The forces generated by the isotropic potentials (eqns. 6.9 and 6.12) also contain components parallel to the rotation axis and thereby restrain motions along the axis of either the whole rotation group (in case of V^{iso}) or within the rotation group (in case of $V^{\text{iso-pf}}$). For cases where unrestrained motion along the axis is preferred, we have implemented a "parallel motion" variant by eliminating all components parallel to the rotation axis for the potential. This is achieved by projecting the distance vectors between reference and actual positions

$$\mathbf{r}_i = \mathbf{\Omega}(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u}) \tag{6.14}$$

onto the plane perpendicular to the rotation vector,

$$\mathbf{r}_i^\perp := \mathbf{r}_i - (\mathbf{r}_i \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}, \tag{6.15}$$

yielding

$$
\begin{aligned}
V^{\text{pm}} &= \frac{k}{2} \sum_{i=1}^{N} w_i (\mathbf{r}_i^\perp)^2 \\
&= \frac{k}{2} \sum_{i=1}^{N} w_i \left\{ \mathbf{\Omega}(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u}) \right. \\
&\quad \left. - \left\{ \left[\mathbf{\Omega}(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u}) \right] \cdot \hat{\mathbf{v}} \right\} \hat{\mathbf{v}} \right\}^2,
\end{aligned} \tag{6.16}
$$

and similarly

$$\mathbf{F}_j^{\text{pm}} = k\, w_j\, \mathbf{r}_j^\perp. \tag{6.17}$$

Pivot-Free Parallel Motion Potential

Replacing in eqn. 6.16 the fixed pivot $\mathbf{u}$ by the center of mass $\mathbf{x}_c$ yields the pivot-free variant of the parallel motion potential. With

$$\mathbf{s}_i = \mathbf{\Omega}(t)(\mathbf{y}_i^0 - \mathbf{y}_c^0) - (\mathbf{x}_i - \mathbf{x}_c) \tag{6.18}$$

the respective potential and forces are

$$V^{\text{pm-pf}} = \frac{k}{2} \sum_{i=1}^{N} w_i (\mathbf{s}_i^\perp)^2, \tag{6.19}$$

$$\mathbf{F}_j^{\text{pm-pf}} = k\, w_j\, \mathbf{s}_j^\perp. \tag{6.20}$$

Radial Motion Potential

In the above variants, the minimum of the rotation potential is either a single point at the reference position $\boldsymbol{y}_i$ (for the isotropic potentials) or a single line through $\boldsymbol{y}_i$ parallel to the rotation axis (for the parallel motion potentials). As a result, radial forces restrict radial motions of the atoms. The two subsequent types of rotation potentials, V^{rm} and V^{rm2}, drastically reduce or even eliminate this effect. The first variant, V^{rm} (Fig. 6.4B), eliminates all force components parallel to the vector connecting the reference atom and the rotation axis,

$$V^{\mathrm{rm}} = \frac{k}{2} \sum_{i=1}^{N} w_i \left[\boldsymbol{p}_i \cdot (\boldsymbol{x}_i - \boldsymbol{u}) \right]^2 , \tag{6.21}$$

with

$$\boldsymbol{p}_i := \frac{\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u})}{\|\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u})\|} . \tag{6.22}$$

This variant depends only on the distance $\boldsymbol{p}_i \cdot (\boldsymbol{x}_i - \boldsymbol{u})$ of atom i from the plane spanned by $\hat{\boldsymbol{v}}$ and $\boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u})$. The resulting force is

$$\mathbf{F}_j^{\mathrm{rm}} = -k\, w_j \left[\boldsymbol{p}_j \cdot (\boldsymbol{x}_j - \boldsymbol{u}) \right] \boldsymbol{p}_j . \tag{6.23}$$

Pivot-Free Radial Motion Potential

Proceeding similar to the pivot-free isotropic potential yields a pivot-free version of the above potential. With

$$\boldsymbol{q}_i := \frac{\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c^0)}{\|\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c^0)\|} , \tag{6.24}$$

the potential and force for the pivot-free variant of the radial motion potential read

$$V^{\mathrm{rm\text{-}pf}} = \frac{k}{2} \sum_{i=1}^{N} w_i \left[\boldsymbol{q}_i \cdot (\boldsymbol{x}_i - \boldsymbol{x}_c) \right]^2 , \tag{6.25}$$

$$\mathbf{F}_j^{\mathrm{rm\text{-}pf}} = -k\, w_j \left[\boldsymbol{q}_j \cdot (\boldsymbol{x}_j - \boldsymbol{x}_c) \right] \boldsymbol{q}_j + k \frac{m_j}{M} \sum_{i=1}^{N} w_i \left[\boldsymbol{q}_i \cdot (\boldsymbol{x}_i - \boldsymbol{x}_c) \right] \boldsymbol{q}_i . \tag{6.26}$$

Radial Motion 2 Alternative Potential

As seen in Fig. 6.4B, the force resulting from V^{rm} still contains a small, second-order radial component. In most cases, this perturbation is tolerable; if not, the following alternative, V^{rm2}, fully eliminates the radial contribution to the force, as depicted in Fig. 6.4C,

$$V^{\mathrm{rm2}} = \frac{k}{2} \sum_{i=1}^{N} w_i \frac{\left[(\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{u})) \cdot \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u}) \right]^2}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{u})\|^2 + \epsilon'} , \tag{6.27}$$

where a small parameter ϵ' has been introduced to avoid singularities. For $\epsilon' = 0\,\mathrm{nm}^2$, the equipotential planes are spanned by $\boldsymbol{x}_i - \boldsymbol{u}$ and $\hat{\boldsymbol{v}}$, yielding a force perpendicular to $\boldsymbol{x}_i - \boldsymbol{u}$, thus not contracting or expanding structural parts that moved away from or toward the rotation axis.

Choosing a small positive ϵ' (*e.g.*, $\epsilon' = 0.01 \, \text{nm}^2$, Fig. 6.4D) in the denominator of eqn. 6.27 yields a well-defined potential and continuous forces also close to the rotation axis, which is not the case for $\epsilon' = 0 \, \text{nm}^2$ (Fig. 6.4C). With

$$\boldsymbol{r}_i \quad := \quad \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{u}) \tag{6.28}$$

$$\boldsymbol{s}_i \quad := \quad \frac{\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{u})}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{u})\|} \equiv \Psi_i \; \hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{u}) \tag{6.29}$$

$$\Psi_i^* \quad := \quad \frac{1}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{u})\|^2 + \epsilon'} \tag{6.30}$$

the force on atom j reads

$$\boldsymbol{F}_j^{\text{rm2}} = -k \left\{ w_j \, (\boldsymbol{s}_j \cdot \boldsymbol{r}_j) \left[\frac{\Psi_j^*}{\Psi_j} \boldsymbol{r}_j - \frac{\Psi_j^{*2}}{\Psi_j^3} (\boldsymbol{s}_j \cdot \boldsymbol{r}_j) \boldsymbol{s}_j \right] \right\} \times \hat{\boldsymbol{v}}. \tag{6.31}$$

Pivot-Free Radial Motion 2 Potential

The pivot-free variant of the above potential is

$$V^{\text{rm2-pf}} = \frac{k}{2} \sum_{i=1}^N w_i \frac{\left[(\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c)) \cdot \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c) \right]^2}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c)\|^2 + \epsilon'} . \tag{6.32}$$

With

$$\boldsymbol{r}_i \quad := \quad \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c) \tag{6.33}$$

$$\boldsymbol{s}_i \quad := \quad \frac{\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c)}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c)\|} \equiv \Psi_i \; \hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c) \tag{6.34}$$

$$\Psi_i^* \quad := \quad \frac{1}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c)\|^2 + \epsilon'} \tag{6.35}$$

the force on atom j reads

$$\begin{aligned}
\boldsymbol{F}_j^{\text{rm2-pf}} \quad = \quad & -k \left\{ w_j \, (\boldsymbol{s}_j \cdot \boldsymbol{r}_j) \left[\frac{\Psi_j^*}{\Psi_j} \boldsymbol{r}_j - \frac{\Psi_j^{*2}}{\Psi_j^3} (\boldsymbol{s}_j \cdot \boldsymbol{r}_j) \boldsymbol{s}_j \right] \right\} \times \hat{\boldsymbol{v}} \\
& + k \frac{m_j}{M} \left\{ \sum_{i=1}^N w_i \, (\boldsymbol{s}_i \cdot \boldsymbol{r}_i) \left[\frac{\Psi_i^*}{\Psi_i} \boldsymbol{r}_i - \frac{\Psi_i^{*2}}{\Psi_i^3} (\boldsymbol{s}_i \cdot \boldsymbol{r}_i) \, \boldsymbol{s}_i \right] \right\} \times \hat{\boldsymbol{v}} .
\end{aligned} \tag{6.36}$$

6.5.2 Flexible Axis Rotation

As sketched in Fig. 6.3A–B, the rigid body behavior of the fixed axis rotation scheme is a drawback for many applications. In particular, deformations of the rotation group are suppressed when the equilibrium atom positions directly depend on the reference positions. To avoid this limitation, eqns. 6.26 and 6.32 will now be generalized towards a "flexible axis" as sketched in Fig. 6.3C. This will be achieved by subdividing the rotation group into a set of equidistant slabs perpendicular to the rotation vector, and by applying a separate rotation potential to each of these slabs. Fig. 6.3C shows the midplanes of the slabs as dotted straight lines and the centers as thick black dots.
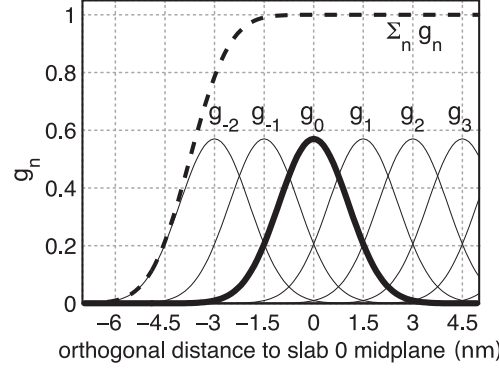
Figure 6.5: Gaussian functions g_n centered at $n\,\Delta x$ for a slab distance $\Delta x = 1.5$ nm and $n \geq -2$. Gaussian function g_0 is highlighted in bold; the dashed line depicts the sum of the shown Gaussian functions.

To avoid discontinuities in the potential and in the forces, we define "soft slabs" by weighing the contributions of each slab n to the total potential function V^{flex} by a Gaussian function

$$g_n(\boldsymbol{x}_i) = \Gamma\,\exp\left(-\frac{\beta_n^2(\boldsymbol{x}_i)}{2\sigma^2}\right), \qquad (6.37)$$

centered at the midplane of the nth slab. Here σ is the width of the Gaussian function, Δx the distance between adjacent slabs, and

$$\beta_n(\boldsymbol{x}_i) := \boldsymbol{x}_i \cdot \hat{\boldsymbol{v}} - n\,\Delta x\,. \qquad (6.38)$$

A most convenient choice is $\sigma = 0.7\Delta x$ and

$$1/\Gamma = \sum_{n \in Z} \exp\left(-\frac{(n-\frac{1}{4})^2}{2\cdot 0.7^2}\right) \approx 1.75464\,,$$

which yields a nearly constant sum, essentially independent of $\boldsymbol{x}_i$ (dashed line in Fig. 6.5), *i.e.*,

$$\sum_{n \in Z} g_n(\boldsymbol{x}_i) = 1 + \epsilon(\boldsymbol{x}_i)\,, \qquad (6.39)$$

with $|\epsilon(\boldsymbol{x}_i)| < 1.3 \cdot 10^{-4}$. This choice also implies that the individual contributions to the force from the slabs add up to unity such that no further normalization is required.

To each slab center $\boldsymbol{x}_c^n$, all atoms contribute by their Gaussian-weighted (optionally also mass-weighted) position vectors $g_n(\boldsymbol{x}_i)\,\boldsymbol{x}_i$. The instantaneous slab centers $\boldsymbol{x}_c^n$ are calculated from the current positions $\boldsymbol{x}_i$,

$$\boldsymbol{x}_c^n = \frac{\sum_{i=1}^N g_n(\boldsymbol{x}_i)\,m_i\,\boldsymbol{x}_i}{\sum_{i=1}^N g_n(\boldsymbol{x}_i)\,m_i}\,, \qquad (6.40)$$

while the reference centers $\boldsymbol{y}_c^n$ are calculated from the reference positions $\boldsymbol{y}_i^0$,

$$\boldsymbol{y}_c^n = \frac{\sum_{i=1}^N g_n(\boldsymbol{y}_i^0)\,m_i\,\boldsymbol{y}_i^0}{\sum_{i=1}^N g_n(\boldsymbol{y}_i^0)\,m_i}\,. \qquad (6.41)$$

Due to the rapid decay of g_n, each slab will essentially involve contributions from atoms located within $\approx 3\Delta x$ from the slab center only.

Flexible Axis Potential

We consider two flexible axis variants. For the first variant, the slab segmentation procedure with Gaussian weighting is applied to the radial motion potential (eqn. 6.26 / Fig. 6.4B), yielding as the contribution of slab n

$$V^n = \frac{k}{2} \sum_{i=1}^{N} w_i \, g_n(\boldsymbol{x}_i) \, [\boldsymbol{q}_i^n \cdot (\boldsymbol{x}_i - \boldsymbol{x}_c^n)]^2 \,,$$

and a total potential function

$$V^{\text{flex}} = \sum_n V^n \,. \tag{6.42}$$

Note that the global center of mass $\boldsymbol{x}_c$ used in eqn. 6.26 is now replaced by $\boldsymbol{x}_c^n$, the center of mass of the slab. With

$$\boldsymbol{q}_i^n \quad := \quad \frac{\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c^n)}{\|\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c^n)\|} \tag{6.43}$$

$$b_i^n \quad := \quad \boldsymbol{q}_i^n \cdot (\boldsymbol{x}_i - \boldsymbol{x}_c^n) \,, \tag{6.44}$$

the resulting force on atom j reads

$$\begin{aligned}
\boldsymbol{F}_j^{\text{flex}} \quad = \quad & -k \, w_j \sum_n g_n(\boldsymbol{x}_j) \, b_j^n \left\{ \boldsymbol{q}_j^n - b_j^n \frac{\beta_n(\boldsymbol{x}_j)}{2\sigma^2} \hat{\boldsymbol{v}} \right\} \\
& + k \, m_j \sum_n \frac{g_n(\boldsymbol{x}_j)}{\sum_h g_n(\boldsymbol{x}_h)} \sum_{i=1}^{N} w_i \, g_n(\boldsymbol{x}_i) \, b_i^n \left\{ \boldsymbol{q}_i^n - \frac{\beta_n(\boldsymbol{x}_j)}{\sigma^2} [\boldsymbol{q}_i^n \cdot (\boldsymbol{x}_j - \boldsymbol{x}_c^n)] \, \hat{\boldsymbol{v}} \right\} \,. \tag{6.45}
\end{aligned}$$

Note that for V^{flex}, as defined, the slabs are fixed in space and so are the reference centers $\boldsymbol{y}_c^n$. If during the simulation the rotation group moves too far in $\boldsymbol{v}$ direction, it may enter a region where – due to the lack of nearby reference positions – no reference slab centers are defined, rendering the potential evaluation impossible. We therefore have included a slightly modified version of this potential that avoids this problem by attaching the midplane of slab $n = 0$ to the center of mass of the rotation group, yielding slabs that move with the rotation group. This is achieved by subtracting the center of mass $\boldsymbol{x}_c$ of the group from the positions,

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{x}_i - \boldsymbol{x}_c \,, \quad \text{and} \quad \tilde{\boldsymbol{y}}_i^0 = \boldsymbol{y}_i^0 - \boldsymbol{y}_c^0 \,, \tag{6.46}$$

such that

$$V^{\text{flex-t}} \quad = \quad \frac{k}{2} \sum_n \sum_{i=1}^{N} w_i \, g_n(\tilde{\boldsymbol{x}}_i) \left[\frac{\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\tilde{\boldsymbol{y}}_i^0 - \tilde{\boldsymbol{y}}_c^n)}{\|\hat{\boldsymbol{v}} \times \boldsymbol{\Omega}(t)(\tilde{\boldsymbol{y}}_i^0 - \tilde{\boldsymbol{y}}_c^n)\|} \cdot (\tilde{\boldsymbol{x}}_i - \tilde{\boldsymbol{x}}_c^n) \right]^2 \,. \tag{6.47}$$

To simplify the force derivation, and for efficiency reasons, we here assume $\boldsymbol{x}_c$ to be constant, and thus $\partial \boldsymbol{x}_c / \partial x = \partial \boldsymbol{x}_c / \partial y = \partial \boldsymbol{x}_c / \partial z = 0$. The resulting force error is small (of order $O(1/N)$ or $O(m_j/M)$ if mass-weighting is applied) and can therefore be tolerated. With this assumption, the forces $\boldsymbol{F}^{\text{flex-t}}$ have the same form as eqn. 6.45.

Flexible Axis 2 Alternative Potential

In this second variant, slab segmentation is applied to V^{rm2} (eqn. 6.32), resulting in a flexible axis potential without radial force contributions (Fig. 6.4C),

$$V^{\mathrm{flex2}} = \frac{k}{2} \sum_{i=1}^{N} \sum_{n} w_i\, g_n(\boldsymbol{x}_i) \frac{[(\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c^n)) \cdot \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c^n)]^2}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c^n)\|^2 + \epsilon'} . \tag{6.48}$$

With

$$\boldsymbol{r}_i^n \quad := \quad \boldsymbol{\Omega}(t)(\boldsymbol{y}_i^0 - \boldsymbol{y}_c^n) \tag{6.49}$$

$$\boldsymbol{s}_i^n \quad := \quad \frac{\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c^n)}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c^n)\|} \equiv \psi_i \ \hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c^n) \tag{6.50}$$

$$\psi_i^* \quad := \quad \frac{1}{\|\hat{\boldsymbol{v}} \times (\boldsymbol{x}_i - \boldsymbol{x}_c^n)\|^2 + \epsilon'} \tag{6.51}$$

$$W_j^n \quad := \quad \frac{g_n(\boldsymbol{x}_j)\, m_j}{\sum_h g_n(\boldsymbol{x}_h)\, m_h} \tag{6.52}$$

$$\boldsymbol{S}^n \quad := \quad \sum_{i=1}^{N} w_i\, g_n(\boldsymbol{x}_i)\, (\boldsymbol{s}_i^n \cdot \boldsymbol{r}_i^n) \left[\frac{\psi_i^*}{\psi_i} \boldsymbol{r}_i^n - \frac{\psi_i^{*2}}{\psi_i^3} (\boldsymbol{s}_i^n \cdot \boldsymbol{r}_i^n)\, \boldsymbol{s}_i^n \right] \tag{6.53}$$

the force on atom j reads

$$\begin{aligned}
\boldsymbol{F}_j^{\mathrm{flex2}} \quad = \quad & -k \left\{ \sum_n w_j\, g_n(\boldsymbol{x}_j)\, (\boldsymbol{s}_j^n \cdot \boldsymbol{r}_j^n) \left[\frac{\psi_j^*}{\psi_j} \boldsymbol{r}_j^n - \frac{\psi_j^{*2}}{\psi_j^3} (\boldsymbol{s}_j^n \cdot \boldsymbol{r}_j^n)\, \boldsymbol{s}_j^n \right] \right\} \times \hat{\boldsymbol{v}} \\
& + k \left\{ \sum_n W_j^n\, \boldsymbol{S}^n \right\} \times \hat{\boldsymbol{v}} - k \left\{ \sum_n W_j^n\, \frac{\beta_n(\boldsymbol{x}_j)}{\sigma^2} \frac{1}{\psi_j}\ \boldsymbol{s}_j^n \cdot \boldsymbol{S}^n \right\} \hat{\boldsymbol{v}} \\
& + \frac{k}{2} \left\{ \sum_n w_j\, g_n(\boldsymbol{x}_j) \frac{\beta_n(\boldsymbol{x}_j)}{\sigma^2} \frac{\psi_j^*}{\psi_j^2} (\boldsymbol{s}_j^n \cdot \boldsymbol{r}_j^n)^2 \right\} \hat{\boldsymbol{v}}.
\end{aligned} \tag{6.54}$$

Applying transformation (6.46) yields a "translation-tolerant" version of the flexible 2 potential, $V^{\mathrm{flex2\text{-}t}}$. Again, assuming that $\partial \boldsymbol{x}_c/\partial x$, $\partial \boldsymbol{x}_c/\partial y$, $\partial \boldsymbol{x}_c/\partial z$ are small, the resulting equations for $V^{\mathrm{flex2\text{-}t}}$ and $\boldsymbol{F}^{\mathrm{flex2\text{-}t}}$ are similar to those of V^{flex2} and $\boldsymbol{F}^{\mathrm{flex2}}$.

6.5.3 Usage

To apply enforced rotation, the particles i that are to be subjected to one of the rotation potentials are defined via index groups `rot_group0`, `rot_group1`, etc., in the `.mdp` input file. The reference positions $\boldsymbol{y}_i^0$ are read from a special `.trr` file provided to `grompp`. If no such file is found, $\boldsymbol{x}_i(t=0)$ are used as reference positions and written to `.trr` such that they can be used for subsequent setups. All parameters of the potentials such as k, ϵ', etc. (Table 6.1) are provided as `.mdp` parameters; `rot_type` selects the type of the potential. The option `rot_massw` allows to choose whether or not to use mass-weighted averaging. For the flexible potentials, a cutoff value $g_n^{\min}$ (typically $g_n^{\min} = 0.001$) makes shure that only significant contributions to V and $\boldsymbol{F}$ are evaluated, *i.e.* terms with $g_n(\boldsymbol{x}) < g_n^{\min}$ are omitted. Table 6.2 summarizes observables that are written to additional output files and which are described below.

Table 6.1: Parameters used by the various rotation potentials. x's indicate which parameter is actually used for a given potential.

parameter			k	$\hat{\boldsymbol{v}}$	$\boldsymbol{u}$	ω	ϵ'	Δx	$g_n^{\min}$
`.mdp` input variable name			k	vec	pivot	rate	eps	slab_dist	min_gauss
unit			$\frac{\text{kJ}}{\text{mol·nm}^2}$	-	nm	°/ps	nm²	nm	-
fixed axis potentials:		eqn.							
isotropic	V^{iso}	(6.9)	X	X	X	X	-	-	-
— pivot-free	$V^{\text{iso-pf}}$	(6.12)	X	X	-	X	-	-	-
parallel motion	V^{pm}	(6.16)	X	X	X	X	-	-	-
— pivot-free	$V^{\text{pm-pf}}$	(6.20)	X	X	-	X	-	-	-
radial motion	V^{rm}	(6.21)	X	X	X	X	-	-	-
— pivot-free	$V^{\text{rm-pf}}$	(6.26)	X	X	-	X	-	-	-
radial motion 2	V^{rm2}	(6.27)	X	X	X	X	X	-	-
— pivot-free	$V^{\text{rm2-pf}}$	(6.32)	X	X	-	X	X	-	-
flexible axis potentials:		eqn.							
flexible	V^{flex}	(6.42)	X	X	-	X	-	X	X
— transl. tol.	$V^{\text{flex-t}}$	(6.47)	X	X	-	X	-	X	X
flexible 2	V^{flex2}	(6.48)	X	X	-	X	X	X	X
— transl. tol.	$V^{\text{flex2-t}}$	-	X	X	-	X	X	X	X

Table 6.2: Quantities recorded in output files during enforced rotation simulations. All slab-wise data is written every `nstsout` steps, other rotation data every `nstrout` steps.

quantity	unit	equation	output file	fixed	flexible
$V(t)$	kJ/mol	see 6.1	`rotation`	X	X
$\theta_{\text{ref}}(t)$	degrees	$\theta_{\text{ref}}(t) = \omega t$	`rotation`	X	X
$\theta_{\text{av}}(t)$	degrees	(6.55)	`rotation`	X	-
$\theta_{\text{fit}}(t), \theta_{\text{fit}}(t,n)$	degrees	(6.57)	`rotangles`	-	X
$\boldsymbol{y}_0(n), \boldsymbol{x}_0(t,n)$	nm	(6.40, 6.41)	`rotslabs`	-	X
$\tau(t)$	kJ/mol	(6.58)	`rotation`	X	-
$\tau(t,n)$	kJ/mol	(6.58)	`rottorque`	-	X

Angle of Rotation Groups: Fixed Axis

For fixed axis rotation, the average angle $\theta_{\mathrm{av}}(t)$ of the group relative to the reference group is determined via the distance-weighted angular deviation of all rotation group atoms from their reference positions,

$$\theta_{\mathrm{av}} = \sum_{i=1}^{N} r_i \, \theta_i \bigg/ \sum_{i=1}^{N} r_i \; . \tag{6.55}$$

Here, r_i is the distance of the reference position to the rotation axis, and the difference angles θ_i are determined from the atomic positions, projected onto a plane perpendicular to the rotation axis through pivot point $\boldsymbol{u}$ (see eqn. 6.15 for the definition of $\perp$),

$$\cos \theta_i = \frac{(\boldsymbol{y}_i - \boldsymbol{u})^\perp \cdot (\boldsymbol{x}_i - \boldsymbol{u})^\perp}{\|(\boldsymbol{y}_i - \boldsymbol{u})^\perp \cdot (\boldsymbol{x}_i - \boldsymbol{u})^\perp\|} \; . \tag{6.56}$$

The sign of θ_{av} is chosen such that $\theta_{\mathrm{av}} > 0$ if the actual structure rotates ahead of the reference.

Angle of Rotation Groups: Flexible Axis

For flexible axis rotation, two outputs are provided, the angle of the entire rotation group, and separate angles for the segments in the slabs. The angle of the entire rotation group is determined by an RMSD fit of $\boldsymbol{x}_i$ to the reference positions $\boldsymbol{y}_i^0$ at $t = 0$, yielding θ_{fit} as the angle by which the reference has to be rotated around $\hat{\boldsymbol{v}}$ for the optimal fit,

$$\mathrm{RMSD}(\boldsymbol{x}_i, \; \boldsymbol{\Omega}(\theta_{\mathrm{fit}})\boldsymbol{y}_i^0) \stackrel{!}{=} \min . \tag{6.57}$$

To determine the local angle for each slab n, both reference and actual positions are weighted with the Gaussian function of slab n, and $\theta_{\mathrm{fit}}(t, n)$ is calculated as in eqn. 6.57) from the Gaussian-weighted positions.

For all angles, the `.mdp` input option `rot_fit_method` controls whether a normal RMSD fit is performed or whether for the fit each position $\boldsymbol{x}_i$ is put at the same distance to the rotation axis as its reference counterpart $\boldsymbol{y}_i^0$. In the latter case, the RMSD measures only angular differences, not radial ones.

Angle Determination by Searching the Energy Minimum

Alternatively, for `rot_fit_method = potential`, the angle of the rotation group is determined as the angle for which the rotation potential energy is minimal. Therefore, the used rotation potential is additionally evaluated for a set of angles around the current reference angle. In this case, the `rotangles.log` output file contains the values of the rotation potential at the chosen set of angles, while `rotation.xvg` lists the angle with minimal potential energy.

Torque

The torque $\boldsymbol{\tau}(t)$ exerted by the rotation potential is calculated for fixed axis rotation via

$$\boldsymbol{\tau}(t) = \sum_{i=1}^{N} \boldsymbol{r}_i(t) \times \boldsymbol{f}_i^\perp(t), \tag{6.58}$$

where $\boldsymbol{r}_i(t)$ is the distance vector from the rotation axis to $\boldsymbol{x}_i(t)$ and $\boldsymbol{f}_i^{\perp}(t)$ is the force component perpendicular to $\boldsymbol{r}_i(t)$ and $\hat{\boldsymbol{v}}$. For flexible axis rotation, torques $\boldsymbol{\tau}_n$ are calculated for each slab using the local rotation axis of the slab and the Gaussian-weighted positions.