Nondimensionalization in MD Simulation

In this document, I would try to elucidate the matter of nondimensionalization utilized in our MD simulation. Intuitively this should be a very straightforward matter, however, as we will see, some minor snags stand, and I shall attempt to make the entire issue clear. Hopefully the completion of this document will make myself and any future reader clear.

1 Equation of Motion

A classical model would be sufficient in the study of our system, so there is essentially only one equation of motion needed in our system, the second law of Newton.

$$m\vec{a} = -\nabla U - \zeta\vec{v} + \vec{f}_{rand} \tag{1}$$

where ζ is the damping factor and f_{rand} is the random force. This is the Langevin equation, and this essentially describes the system in whole.

However, to fully describe the system, we need to specify the force fields and potentials. The potentials can be generally divided into two categories, the bonding potentials between adjacent polymer units, and non-bonding potentials.

$$U = U_{bond} + U_{non-bond} \tag{2}$$

The bonding potentials make up of the vibration potential and the bending potential

$$\begin{aligned}
U_{bond} &= U_{vib} + U_{bend} && (3)\\
U_{vib} &= \frac{k_{vib}k_BT}{2\sigma_M^2}\sum\left(|\vec{r}_i - \vec{r}_{i+1} - \sigma_M\right)^2 && (4)\\
U_{bend} &= k_{bend}k_BT\sum\left(1 - \frac{(\vec{r}_{i-1} - \vec{r}_i)(\vec{r}_i - \vec{r}_i - 1)}{|\vec{r}_{i-1} - \vec{r}_i||\vec{r}_i - \vec{r}_i - 1|}\right) && (5)
\end{aligned}$$

and the non-bonding potentials include the excluded-volume (ev) potential, the wall interaction potential and potential due to external conservative force fields such as electrostatic potential

$$\begin{aligned}
U_{non-bond} &= U_{ev} + U_{wall} + U_{ext} && (6)\\
U_{ev} &= \sum_{i=1,j>i+1}^{i=j=N} 4\epsilon_m k_BT\left[\left(\frac{\sigma_m}{r_{ij}}\right)^{12} - \left(\frac{\sigma_m}{r_{ij}}\right)^{6} + \frac{1}{4}\right] && (7)
\end{aligned}$$

the excluded volume potential adopted in current study is the pair truncated Lennard-Jones potential. The wall potential is also a modified version of the LJ potential, however, for brevity reasons, it is not shown here because the nondimensionalization process is essentially identical to the excluded volume potential.

Now, we can observe that intrinsically from our construction of the vibration potential and bending potential, the incorporation of the factor k_BT and σ_M imply the nondimensionalization with these factors. As we will soon see, this is indeed the case.

2 Selecting Parameters for Nondimensionalization

Now, we try to find a simple set of parameters to nondimensionalize the set of equation above. From our construction, it is clear that we should opt for k_BT as the characteristic energy scale, and σ_M as the characteristic length scale. Now we try this, and the equations become

$$U^*_{vib} = \frac{k_{vib}}{2}\sum\left(|\vec{r}^*_i - \vec{r}^*_{i+1} - 1\right)^2 \tag{8}$$

$$U^*_{bend} = k_{bend}\sum\left(1 - \frac{(\vec{r}^*_{i-1} - \vec{r}^*_i)(\vec{r}^*_i - \vec{r}^*_i - 1)}{|\vec{r}^*_{i-1} - \vec{r}^*_i||\vec{r}^*_i - \vec{r}^*_i - 1|}\right) \tag{9}$$

$$U^*_{ev} = \sum_{i=1,j>i+1}^{i=j=N} 4\epsilon_m\left(r^{*-12}_{ij} - r^{*-6}_{ij} + \frac{1}{4}\right) \tag{10}$$

where

$$U = U^* k_B T \tag{11}$$

$$r = r^* \sigma_M \tag{12}$$

We see that the potentials are successfully transformed into nondimensional equations.

Now we tackle the equation of motion. Intuitively, we shall choose a character timescale to proceed. But for reasons that would be demonstrated in following text, we choose instead the damping factor ζ as the characteristic dimension to work with. First, we identify with Stokes' law

$$F_d = 6\pi\eta R v = 6\pi\eta\frac{\sigma_M}{2}v = \zeta v \tag{13}$$

so we see that the damping coefficient in terms of dimension is

$$\zeta \quad [=] \quad \eta\sigma_M = [M][T]^{-1}[L]^{-1} * [L] = [M][T]^{-1} \tag{14}$$

where $[=]$ denotes the equivalence of the physical dimension, and M, T, L, F stands for mass, time, length and force respectively. We now dismantle the dimensions we selected earlier into other dimensions.

$$k_BT \quad [=] \quad [E] = [M][L][T]^{-2} * [L] = [M][L]^2[T]^{-2} \tag{15}$$

$$\sigma_M \quad [=] \quad [L] \tag{16}$$

We could now use these to form other dimensions!

$$[T] = ([M][T]^{-1}) * ([M][L]^2[T]^{-2})^{-1} * ([L])^2 \tag{17}$$

$$[M] = ([M][T]^{-1})^2 * ([M][L]^2[T]^{-2})^{-1} * ([L])^2 \tag{18}$$

$$[F] = [E][L]^{-1} \tag{19}$$

Therefore it follows that time, mass and force should be nondimensionalized with

$$t = \tau t^* = \left(\frac{\zeta\sigma_M^2}{k_BT}\right)t^* \tag{20}$$

$$m = \Lambda m^* = \left(\frac{\zeta^2\sigma_M^2}{k_BT}\right)m^* \tag{21}$$

$$f = \Phi f^* = \left(\frac{k_BT}{\sigma_M}\right)f^* \tag{22}$$

To verify this, let us try to nondimensionalize the equation of motion directly.

$$
\begin{aligned}
& m\vec{a} = -\nabla U - \zeta\vec{v} + \vec{f}_{rand} && (23)\\
\Rightarrow\ & m\vec{a}\left(\frac{\sigma_M}{k_BT}\right) = -\nabla U\left(\frac{\sigma_M}{k_BT}\right) - \zeta\vec{v}\left(\frac{\sigma_M}{k_BT}\right) + \vec{f}_{rand}\left(\frac{\sigma_M}{k_BT}\right) && (24)\\
\Rightarrow\ & m\vec{a}\left(\frac{k_BT}{\sigma_M^2\zeta^2}\right)\left(\frac{\zeta\sigma_M^2}{k_BT}\right)^2\left(\frac{1}{\sigma_M}\right) = -\nabla U\left(\frac{1}{k_BT}\right)(\sigma_M) - \vec{v}\left(\frac{\zeta\sigma_M^2}{k_BT}\right)\left(\frac{1}{\sigma_M}\right) + \vec{f}^*_{rand} &&\\
\Rightarrow\ & m^*\vec{a}^* = -\nabla^* U^* - \vec{v}^* + \vec{f}^*_{rand} && (25)
\end{aligned}
$$

where

$$
\begin{aligned}
\vec{a}^* &= \frac{\partial^2\vec{r}^*}{\partial t^{*2}}\\
\vec{v}^* &= \frac{\partial\vec{r}^*}{\partial t^*}\\
\nabla^* &= \hat{e}^*_i\frac{\partial}{\partial x^*_i} \qquad (26)
\end{aligned}
$$

Other important quantities can also be nondimensionalized following the same procedure.

- Dynamic viscosity of solvent: $\eta = (\zeta/3\pi\sigma_M) \Rightarrow \eta^* = \eta\left(\zeta/\sigma_M\right)^{-1} = 1/3\pi$ (this follows from the Stokes' equation)
- Diffusivity of monomer: $D_m = (k_BT/3\pi\eta\sigma_M) \Rightarrow D^*_m = D_m(k_BT/\zeta)^{-1} = 1$ (this follows from the Stokes-Einstein equation)
- Diffusion timescale of a momoner: $\tau_m = (\sigma_M^2/D_m) \Rightarrow \tau^*_m = \tau_m(\zeta\sigma_M^2/k_BT)^{-1} = 1$
- Mass density of a monomer: $\rho_m = (6m/\pi\sigma_M^3) \Rightarrow \rho^*_m = \rho_m(\Lambda/\sigma_M^3)^{-1} = (6m^*/\pi)$

3 Interpreting the Characteristic Quantities

Now we have made our system nondimensional, we need to justify the physical meaning of the quantities. The quantity k_BT is very straightforward to understand. It is a measure of the energy scale for thermal fluctuation. However, other quantities do not follow so intuitively.

We have implicitly suggested that the quantity σ_M to be the diameter of monomers. However, the monomer here does not refer to the monomer molecules that actually constitute the polymer molecules. Rather, by monomer here we actually mean the hypothetical ”soft” beads that are connected together to form a bead-spring chain. The model chain consist of the characteristic of a Gaussian chain (quadratic potential between adjacent nodes), long range monomer interactions (LJ excluded volume effect), and the discrete worm-like chain (potential depends on bond angle). Hence, the quantity σ_M cannot be construed as the bond length of constituting monomers, but rather they are the diameter soft beads.

That being said, we still lack the physical correspondence to measurable physical quantities. What does σ_M actually correspond to? To answer this, we must refer to simulations.

It was found that for simulations with $k_{bend} = 5$, results yield a persistence length $l_p = 5$; while it was also found that for $k_{bend} = 25$, $l_p = 25$. Note that the choice of k_{vib} here should be irrelevant here, since persistence length is a quantity relating to the torsion and bending of a polymer chain. However, for finite simulations, it is expected that for softer chains (small k_{vib}) the fluctuation might be larger. Nonetheless, the average should be identical. The strength of the excluded volume effect is also expected not to affect the result, since the persistence length is a property relevant only to relatively short range interaction. The mass of monomers and the damping factor are also irrelevant, since they contribute to the dynamics only.

Now, l_p is a physical quantity directly measurable from experiments. For dsDNA in a buffer solution, l_p is found to be 50 nm. Hence for $k_{bend} = 5$, this corresponds to $\sigma_M = 10$ nm. From our model, we can also read that the with of DNA chain is also 10 nm (because the diameter of the beads correspond to the with of the chain). With the fact that dsDNA is actually 2 nm wide, we can calculate the Debye length of DNA to be

$$\frac{1}{\kappa} = \frac{10-2}{2} = 4 \text{ (nm)} \tag{27}$$

This correspond to the case of moderate to low ionic strength. In comparison, the case of $k_b = 25$ gives $\sigma_M = 2$ nm ($\because l_p = 25\sigma_M = 50$ nm), and it gives a Debye length on Angstrom scale. This corresponds to the case of very strong ionic strength. Therefore, we could see that the selection of the parameter k_{bend} can be tuned to match the l_p to $1/\kappa$ ratio, which is a experimental factor determined by the composition of the solution. It can be shown that for the range between $5 < k_{bend} < 25$, the correlation between σ_M and k_{bend} is almost linear.

Next we try to determine the value of the remaining parameter. While we have used the damping coefficient ζ as the basic nondimensionalization parameter, the actual physical quantity which we can correlate to is the dynamic viscosity η of the solvent. For pure water, $\eta = 0.89\times10^{-3}$ Pa·s, and since $\zeta = 3\pi\eta\sigma_M$, we have $\zeta = 3\pi\cdot0.89\times10^{-3}\cdot10^{-8} = 8.388\times10^{-11}$ kg/s (for $\sigma_M = 10$ nm).

Now, we can evaluate every parameter in SI units. For the two different k_{bend}, at $T = 25°C$, we have

k_{bend}	5	10	25
σ_M (m)	1×10^{-8}	5×10^{-9}	2×10^{-9}
k_BT (J)	4.11×10^{-21}	4.11×10^{-21}	4.11×10^{-21}
ζ (kg/s)	8.39×10^{-11}	4.71×10^{-11}	1.68×10^{-11}
η (Pa·s)	8.9×10^{-4}	8.9×10^{-4}	8.9×10^{-4}
τ (s)	2.04×10^{-6}	2.87×10^{-7}	1.63×10^{-8}
τ_m (s)	2.04×10^{-6}	2.87×10^{-7}	1.63×10^{-8}
Φ (N)	4.11×10^{-13}	8.22×10^{-13}	2.06×10^{-12}
Λ (kg)	1.71×10^{-16}	1.35×10^{-17}	2.74×10^{-19}
D_m (m^2/s)	4.90×10^{-11}	8.72×10^{-11}	2.45×10^{-10}
ρ_m/m^* (kg/m^3)	3.27×10^{8}	2.06×10^{8}	6.54×10^{7}

In the simulation, iteration timestep is set to $dt = 0.0005$, while $\sigma_M = 5$ nm, $\eta = 1$ cP and $T = 25°C$. This corresponds to $dt = 143$ ps.

4 Choosing Monomer Mass

While we have successfully transformed our simulation into a nondimensional form, with the correlation between actual system clearly demonstrated, there remained one crucial parameter value to be specified, m^*, the nondimensional mass of a monomer. Physically, this quantity should be estimated with σ_M by counting the number of atoms inside the effective volume of a model bead. Another way of estimating is through the estimation of mass density. From the table, we found that the monomer mass density is around $10^8 \cdot m^*$ kg/m^3. Physically, we expect the monomers to be of similar density as the solvent, i.e. water, $m^* \sim 10^{-5}$.

At this point, it seems that we could conclude our discussion. However, one obstacle exists. In our simulation, we rely on numerical method to solve the differential equations. A problem with numerical solution of DE is that it might fail due to instability. Let us work with the simplest method, the forward Euler (FE) method. In our system, the dynamic is essentially overdamped. Hence, in short time scales, we may divide the velocity into two parts, the time-independent drift velocity and the time-varying relaxation velocity (we directly employ the nondimensional version)

$$v^*(t) = v^*_{drift} + v^*_{relax}(t) \tag{28}$$

The equation of motion suitable here is the Langevin equation

$$m^* \frac{dv^*(t)}{dt^*} = f^* + \zeta(t) - v^*(t) \tag{29}$$

Ignoring the random fluctuation (or equivalently, taking the ensemble average), we have

$$m^* \frac{d}{dt^*} \left(v^*_{drift} + v^*_{relax}(t) \right) = f^* - v^*_{drift} - v^*_{relax}(t) \tag{30}$$

At large time, v_{relax} decays and the drift velocity follows the relation

$$v^*_{drift} = f^* \tag{31}$$

and the relaxing velocity follow the DE

$$m^* \frac{dv^*_{relax}(t)}{dt^*} = -v^*_{relax}(t) \tag{32}$$

For ODE in the form $y' = ky$, the stability criterion is

$$|hk + 1| \leq 1 \tag{33}$$

Theorem 1. *The stability criterion for forward Euler method for ODE $y' = ky$ is $|hk + 1| \leq 1$, where h is the integration step size.*

Proof. Since $y_{n+1} = y_n + hy' = y_n + hky_n = (1 + hk)y_n = Gy_n$, $y_{N+1} = G^N y_1$. For y_{N+1} to converge, $|G| \leq 1$, hence $|hk + 1| \leq 1$. □

In our case, we have criterion $|1 - dt^*/m^*$, which translates as $0 \leq dt^*/m^* \leq 2$. The first inequality is trivial, and the latter gives

$$dt^* \leq 2m^* \tag{34}$$

As a consequence, we found that in order for the numeric integration to be stable, our step size is limited by the value of m^*. Now, if we select $m^* \sim 10^{-5}$, the computation effort required to for time scales of our interest would be enormous. This is practically infeasible.

To avoid this, we choose $m^* = 1$. Now the conflict emerges, how do we justify this unphysical selection of monomer mass? The key lies again at the time scale that we are interested in. For a DNA translocation process, typical translocation time lies at 75-400 μs, which is 250-1300 in simulation time. Therefore the time scale for the dynamics of our interest is about 3-10, which is large compared to the relaxation time scale. To verify this, we observe that for relaxation time scale $m_1^* \sim 10^{-5}$ and $m_2^* \sim 1$, $\exp(-3/10^{-5}) = 0$, whereas $\exp(-3) \sim 0.05$, the results were comparable, justifying our arbitrary selection of m^*. In short, the justification for choosing $m^* = 1$ is that the time scale of our interest is also very large compared to 1, therefore in the overdamped limit, we may readjust the relaxation without affecting the dynamics that we are interested in.