Theory behind String Method

Kaihang Shi

I. Restrained MD and Free Energy Gradient (1/29/19)

Suppose we introduce one collective variable, $\theta(r^N)$. The free energy associated with this collective variable is a function depending on z, and the free energy is defined as

$$\begin{aligned} F(z) &= -k_B T \ln Q \\ &= -k_B T \ln \int \dots \int \exp(-\beta U(r^N))\delta(\theta(r^N) - z) dr^N \end{aligned} \tag{1}$$

where $\beta = 1/(k_B T)$; $U(r^N)$ is the potential energy depending on all the cartesian coordinates. Variable z is actually the independent variable in the collective variable space. Now, we rewrite the Dirac delta function in terms of the Gaussian distribution,

$$\delta(\theta(r^N) - z) = \lim_{\sigma \to 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\theta(r^N) - z}{\sigma}\right)^2} \tag{2}$$

If we rewrite the above equation to the form in consistent with the Boltzmann factor, by changing of variable using $\sigma = \sqrt{1/k\beta}$, Eq. (2) can be rearranged to

$$\delta(\theta(r^N) - z) = \lim_{k \to +\infty} \frac{\sqrt{k\beta}}{\sqrt{2\pi}} e^{-\frac{k\beta\left(\theta(r^N) - z\right)^2}{2}} \tag{3}$$

Plugging Eq. (3) into Eq. (1),

$$F(z) = -\lim_{k \to +\infty} k_B T \ln\left[\frac{\sqrt{k\beta}}{\sqrt{2\pi}} \int \dots \int \exp\left[-\beta\left(U(r^N) + \frac{k\left(\theta(r^N) - z\right)^2}{2}\right)\right] dr^N\right] \tag{4}$$

The derivative of free energy in terms of its variable z gives

$$
\begin{aligned}
\frac{\partial F(z)}{\partial z} &= -\lim_{k\to+\infty} k_B T \ln\left[\frac{\sqrt{k\beta}}{\sqrt{2\pi}}\int\ldots\int \exp\left[-\beta\left(U(r^N)+\frac{k\left(\theta(r^N)-z\right)^2}{2}\right)\right]dr^N\right] \\
&= \frac{\partial}{\partial z}\left[-\lim_{k\to+\infty} k_B T \ln\frac{\sqrt{k\beta}}{\sqrt{2\pi}}\right]-\frac{\partial}{\partial z}\left[\lim_{k\to+\infty} k_B T \ln\left[\int\ldots\int \exp\left[-\beta\left(U(r^N)+\frac{k\left(\theta(r^N)-z\right)^2}{2}\right)\right]dr^N\right]\right] \\
&= \lim_{k\to+\infty}\frac{\int\ldots\int k\left(\theta(r^N)-z\right)\exp\left[-\beta\left(U(r^N)+\frac{k\left(\theta(r^N)-z\right)^2}{2}\right)\right]dr^N}{\int\ldots\int \exp\left[-\beta\left(U(r^N)+\frac{k\left(\theta(r^N)-z\right)^2}{2}\right)\right]dr^N} \\
&= \lim_{k\to+\infty}\left\langle k\left(\theta(r^N)-z\right)\right\rangle_{restrain}
\end{aligned}
\tag{5}
$$

Where k is the harmonic force constant corresponding to the variance in the Gaussian distribution. In practice, we want a low force constant to accelerate the equilibrium of replicas during MD. We perform a restrained molecular dynamics and calculate the harmonic force. The force constant is usually chosen as the minimum value that can ensure the independence of restrained force on the force constant.

II. Calculation of the potential of mean force along the minimum free energy path (MFEP) 1/31/2019

First, lets say we discretized the string into $N+1$ representative images,

$$z_n = z(n/N) \qquad \text{n=0,1,2,...N} \tag{6}$$

Upon equilibrium of the string evolution, we can calculate the free energy difference between $F(z_n)$ and the initial point, $F(z_0)$, by doing line integral along the MFEP. For the line interal in a vector field, we have

$$
\begin{aligned}
F(z_n)-F(z_0) &= \int_C \nabla F d\mathbf{z} = \int \nabla F \mathbf{z}' d\alpha \\
&= \sum_{m=1,\Delta\alpha\to 0}^{n} \frac{1}{2}\left(-\nabla F_{m-1}\frac{\mathbf{z}_{m-1}-\mathbf{z}_m}{\Delta\alpha}+\nabla F_m \frac{\mathbf{z}_m-\mathbf{z}_{m-1}}{\Delta\alpha}\right)\Delta\alpha \\
&= \sum_{m=1}^{n}\frac{\left(\nabla F_m+\nabla F_{m-1}\right)}{2}\left(\mathbf{z}_m-\mathbf{z}_{m-1}\right)
\end{aligned}
\tag{7}
$$

Where C is the contour along MFEP; α is the reaction coordinate (or in another words, dummy variable). The Last two steps have used the trapezoidal numerical integration rule. The final results is a summation of a dot product of the free energy gradient and the collective variable vector. Because both the free energy gradient can collective variable vector is pointing along the MFEP, the dot product yields the free energy values.