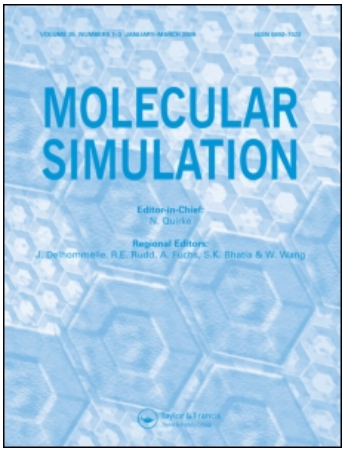


**Molecular Simulation**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

# An Adiabatic Molecular Dynamics Method for the Calculation of Free Energy Profiles

Lula Rosso[a]; Mark E. Tuckerman[a]
[a] Department of Chemistry, New York University, New York, NY

Online publication date: 26 October 2010

**To cite this Article** Rosso, Lula and Tuckerman, Mark E.(2002) 'An Adiabatic Molecular Dynamics Method for the Calculation of Free Energy Profiles', Molecular Simulation, 28: 1, 91 — 112
**To link to this Article: DOI:** 10.1080/08927020211977
**URL:** http://dx.doi.org/10.1080/08927020211977

*Molecular Simulation*, 2002, Vol. 28(1–2), pp. 91 – 112

Taylor & Francis
Taylor & Francis Group

# AN ADIABATIC MOLECULAR DYNAMICS METHOD FOR THE CALCULATION OF FREE ENERGY PROFILES

LULA ROSSO[a] and MARK E. TUCKERMAN[b,*]

[a]*Department of Chemistry, New York University, New York, NY 10003;*
[b]*Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, New York, NY 10003*



*(Received September 2000; accepted January 2001)*

A new molecular dynamics method for calculating free energy profiles for rare events is presented. The new method is based on the creation of an adiabatic separation between the reaction coordinate subspace and the remaining degrees of freedom within a molecular dynamics run. This is achieved by associating with the reaction coordinate(s) a high temperature and large mass, thereby allowing the activated process to occur while permitting the remaining degrees of freedom to respond adiabatically. In this limit, by applying a formal multiple time scale Liouville operator factorization, it can be rigorously shown that the free energy profiles are obtained directly from the probability distribution of the reaction coordinate subspace and, therefore, require no postprocessing of the output data. The new method is applied to a variety of model problems and its performance tested against free energy calculations using the "bluemoon ensemble" approach. The comparison shows that free energy profiles can be calculated with greater ease and efficiency using the new method.

*Keywords*: Molecular dynamics; Energy profiles; AFED; Isomerization; Harmonic oscillator

## 1. INTRODUCTION

One of the most important quantities in thermodynamics is the reversible work needed to change the thermodynamic state of a system. Under certain conditions, this quantity will be independent of the path taken between the initial and final states and will, therefore, be related to the free energy difference between these two states. As a result, considerable effort has been

*Corresponding author.

ISSN: 0892-7022  DOI: 10.1080/08927020290004403

invested in the development of methods to compute such free energy differences (see, *e.g.*, Ref. [1] for a review and comparison). Of particular importance is the free energy profile along a reaction path characterized by a reaction coordinate $q$. Not only does the free energy profile provide a thermodynamic picture along the reaction path, but it also permits determination of activation energies and associated rate constants *via* classical or quantum transition state theory. Statistical mechanics provides a means whereby such free energy profiles can be determined directly from ensemble averages. In the case of a classical $N$-particle system at temperature, $T$, the reaction coordinate, $q$, is expressible as a function $q = q(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ of the $N$ Cartesian position vectors $\{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$ of the $N$ particles. The free energy profile, $F(q')$, is then defined by

$$F(q') = -\frac{1}{\beta} \ln P(q') \tag{1}$$

where

$$P(q') = \langle \delta(q(\mathbf{r}_1, \ldots, \mathbf{r}_N) - q') \rangle \tag{2}$$

is the probability density for the reaction coordinate to take on the value $q'$ and $\beta = 1/kT$.

In principle, the probability distribution in Eq. (2) can be computed directly from a molecular dynamics (MD) or Monte Carlo simulation. However, if the reaction path described by $q$ corresponds to a rare event with a high activation energy, then the use of direct simulation techniques is infeasible due to the presence of very low probability regions in the configuration space. Consequently special techniques such as umbrella sampling [2–4] and the so called "bluemoon ensemble" approach [5, 6] have been developed. In these approaches, the reaction coordinate is restrained to lie within certain windows or rigidly fixed at certain values and individual simulations in each window or at each fixed value are performed. The free energy profile is then reconstructed *a posteriori via* weighted histogram (see, *e.g.*, Ref. [7]) or thermodynamic integration techniques. Despite their success, the necessity of performing numerous individual simulations is nevertheless computationally intensive. In addition, the need to change the reaction coordinate by hand in these methods may involve many nontrivial changes in the other degrees of freedom in a system which require long equilibration times to relax or may, itself, be nontrivial depending on the complexity of the reaction coordinate. Finally, both the umbrella and bluemoon approaches require nontrivial postprocessing of the output data.

In this paper, an alternative approach to the calculation of free energy profiles along reaction paths is presented. The new method is based on the creation of a dynamical adiabatic separation between the reaction coordinate and remaining degrees of freedom. In particular, a dynamical scheme is constructed in which the reaction coordinate evolves slowly relative to the other degrees of freedom and is simultaneously maintained at a high temperature. The latter condition, which has also been employed on entire molecules by other authors to enhance sampling of configuration space [8, 9] and which has been used to perform approximate quantum dynamical simulations [10, 11], ensures that all activation barriers along the reaction path can be easily crossed and can be enforced by coupling the reaction coordinate to its own heat bath or thermostat. The former condition permits the remaining degrees of freedom to fully relax in response to the motion of the reaction coordinate and, thereby, sample a large portion of their configuration space as the reaction coordinate slowly evolves. A careful analysis of the resulting dynamics reveals that the free energy profile will then be given by Eq. (1) with $\beta$ replaced by $\beta_q = 1/kT_q$, where $T_q$ is the temperature of the reaction coordinate (see Section 2). It is then clear that the use of such an approach eliminates all postprocessing of the simulation data. It is also found that the new adiabatic free energy dynamics (AFED) method allows the free energy profile to be determined with greater efficiency than constrained/restrained methods such as the bluemoon and umbrella sampling schemes.



This paper is organized as follows. In Section 2, the adiabatic dynamics method is analyzed in detail and shown to yield the free energy profile directly from Eq. (1) with $\beta$ replaced by $\beta_q$. The analysis is based on the use of the Liouville operator technique and a formally exact multiple time scale breakup of the adiabatic classical propagator. In Section 3, several model problems are examined using the new approach. In these examples, direct comparison to analytical solutions (when available) and the bluemoon ensemble method are made. Conclusions and prospects for extending this work to more complex situations, other types of free energy calculations, and quantum free energy profiles are discussed in Section 4.

## 2. THEORETICAL ANALYSIS OF THE ADIABATIC DYNAMICS APPROACH

In order to illustrate the method, we consider a simple model system, which serves as a paradigm for the general problem of determining free energy

profiles. The model consists of two degrees of freedom, namely, a reaction coordinate $x$ with mass $m_x$ and an additional degree of freedom $y$ with mass $m_y$. The two degrees of freedom are coupled through a potential $V(x, y)$, and the Hamiltonian of the system is taken to be of the form:

$$H = \frac{p_x^2}{2m_x} + \frac{p_y^2}{2m_y} + V(x, y) \tag{3}$$

The simple analysis based on Eq. (3) neglects the fact that, in generalized coordinates, the kinetic energy will actually involve a (generally) coordinate-dependent mass metric tensor. A more general analysis and procedure for treating $N$-particle systems in terms of generalized coordinates is given in Appendix A.

If $\omega_x$ and $\omega_y$ are characteristic frequencies of the $x$ and $y$ motion, then an adiabatic separation between $x$ and $y$ is achieved by requiring that $\omega_x \ll \omega_y$. This can be realized dynamically by choosing the masses such that $m_x \gg m_y$. In addition, temperatures $T_x$ and $T_y$ associated with the two degrees of freedom are introduced such that $T_x \gg T_y$, as noted above. In the proceeding discussion, a detailed analysis of the dynamics and the phase space it generates will be carried out. The goal of this analysis is to show that, under the usual assumptions of ergodicity, the resulting phase space distribution function of the reaction coordinate, $x$, leads to a free energy profile given by Eq. (1) with $\beta$ replaced by $\beta_x = 1/kT_x$.



In order to maintain the two temperatures, $T_x$ and $T_y$ in the system, the dynamics generated by Eq. (3) must be supplemented by coupling the variables $x$ and $y$ to independent heat baths or thermostats. Thus, the phase space evolution of the system is governed by a Liouville operator of the form

$$iL = \frac{p_x}{m_x}\frac{\partial}{\partial x} + \frac{p_y}{m_y}\frac{\partial}{\partial y} + F_x\frac{\partial}{\partial p_x} + F_y\frac{\partial}{\partial p_y} + iL_{\text{therm}}^{(x)}(T_x) + iL_{\text{therm}}^{(y)}(T_y) \tag{4}$$

where $F_x = -\partial V/\partial_x$, $F_y = -\partial V/\partial y$ and $iL_{\text{therm}}^{(x)}(T_x)$ and $iL_{\text{therm}}^{(y)}(T_y)$ are dynamical thermostates, *e.g.*, the Nosé-Hoover chain (NHC) [12] or the recently introduced generalized Gaussian moment thermostat (GGMT) [13], which maintain $x$ at temperature $T_x$ and $y$ at temperature $T_y$. The explicit forms of the thermostat operators are not important for the present discussion, however, the interested reader is referred to Refs. [17] and [13] for the detailed expressions of these operators. For the present discussion, it is sufficient to know that they act on the momenta $p_x$ and $p_y$ in such a way as to control the kinetic energy fluctuations and generate a Maxwell-Boltzmann distribution in these variables at the required temperatures and that they

involve a set of additional or "extended" phase space variables, denoted generally by $\Gamma_{\text{therm},x}$ and $\Gamma_{\text{therm},y}$. It is, again, stressed that two separate thermostats are needed in order to maintain the two separate temperatures $T_x$ and $T_y$.

The time evolution of the full phase space $\Gamma = \{x, y, p_x, p_y, \Gamma_{\text{therm},x}, \Gamma_{\text{therm},y}\}$, starting from an initial condition $\Gamma(0)$, is formally given by

$$\Gamma(t) = e^{iLt}\Gamma(0) \tag{5}$$

where $\exp(iLt)$ is the classical propagator. Consider the evolution of the system over a time interval $\Delta t$ characteristic of the $x$ motion. In order to analyze the dynamics over such an interval, it is useful to define "reference system" Liouville operators:

$$\begin{aligned} iL_{\text{ref}}^{(x)}(T_x) &= \frac{p_x}{m_x}\frac{\partial}{\partial x} + iL_{\text{therm}}^{(x)}(T_x) \\ iL_{\text{ref}}^{(y)}(T_y) &= \frac{p_y}{m_y}\frac{\partial}{\partial y} + iL_{\text{therm}}^{(y)}(T_y) \end{aligned} \tag{6}$$

and express the total Liouville operator as

$$iL = iL^{(y)} + iL_{\text{ref}}^{(x)} + F_x\frac{\partial}{\partial p_x} \tag{7}$$

where

$$iL^{(y)} = iL_{\text{ref}}^{(y)} + F_y\frac{\partial}{\partial p_y} + F_x\frac{\partial}{\partial p_x} \tag{8}$$

Using the Trotter theorem, a reversible, symplectic factorization of the classical propagator is constructed according to

$$\exp(iL\Delta t) = \exp\left(iL^{(y)}\frac{\Delta t}{2}\right)\exp\left(iL_{\text{ref}}^{(x)}\Delta t\right)\exp\left(iL^{(y)}\frac{\Delta t}{2}\right) \tag{9}$$

In Eq. (9), note that the $y$-propagator, $\exp(iL^{(y)}\,\Delta t/2)$, is left intact. The factorization scheme in Eq. (9) allows the $y$-propagator to be evaluated formally exactly, again, using the Trotter theorem.

$$\begin{aligned} \exp\left(iL^{(y)}\frac{\Delta t}{2}\right) = \lim_{n\to\infty}\Bigg[&\exp\left(\frac{\Delta t}{4n}F_x\frac{\partial}{\partial p_x}\right) \\ &\times \exp\left(\frac{\Delta t}{4n}F_y\frac{\partial}{\partial p_y}\right)\exp\left(iL_{\text{ref}}^{(y)}\frac{\Delta t}{2n}\right)\exp\left(\frac{\Delta t}{4n}F_y\frac{\partial}{\partial p_y}\right) \\ &\times \exp\left(\frac{\Delta t}{4n}F_x\frac{\partial}{\partial p_x}\right)\Bigg]^n \end{aligned} \tag{10}$$

When Eq. (10) is substituted into Eq. (9), the result is the formally exact analog of a multiple time scale factorization such as is discussed in Ref. [14]. The formally exact evaluation of the $y$-propagator is necessary since, under the adiabatic conditions of the problem, the time interval, $\Delta t$, is very long compared to the time scale on which $y$ evolves. Combining Eqs. (9) and (10) and acting with the resulting operator on the full phase space vector $\Gamma(0)$ yields the following evolution of the physical variables [11]:

$$\begin{aligned}
x(\Delta t) &= x_{\mathrm{ref}}[x(0), \dot{x}(\Delta t/2), \Gamma_x(0); \Delta t] \\
\dot{x}(\Delta t) &= \dot{x}_{\mathrm{ref}}[x(0), \dot{x}(\Delta t/2), \Gamma_x(0); \Delta t] \\
&\quad + \frac{1}{m_x}\int_{\Delta t/2}^{\Delta t} dt' F_x[y_{\mathrm{adb}}(y(\Delta t/2), \dot{y}(\Delta t/2), \\
&\qquad \Gamma_y(\Delta t/2), x(\Delta t); t'), x(\Delta t)] \\
\dot{x}(\Delta t/2) &= \dot{x}(0) + \frac{1}{m_x}\int_0^{\Delta t/2} dt' F_x[y_{\mathrm{adb}}(y(0), \dot{y}(0), \Gamma_y(0), x(0); t'), x(0)] \\
y(\Delta t/2) &= y_{\mathrm{adb}}(y(0), \dot{y}(0), \Gamma_y(0), x(0); \Delta t/2) \\
\dot{y}(\Delta t/2) &= \dot{y}_{\mathrm{adb}}(y(0), \dot{y}(0), \Gamma_y(0), x(0); \Delta t/2) \\
y(\Delta t) &= y_{\mathrm{adb}}(y(\Delta t/2), \dot{y}(\Delta t/2), \Gamma_y(\Delta t/2), x(\Delta t); \Delta t) \\
\dot{y}(\Delta t) &= \dot{y}_{\mathrm{adb}}(y(\Delta t/2), \dot{y}(\Delta t/2), \Gamma_y(\Delta t/2), x(\Delta t); \Delta t)
\end{aligned} \tag{11}$$

In Eq. (11), $x_{\mathrm{ref}}[x', \dot{x}', \Gamma_x'; t']$ indicates the evolution of $x$ under the action of the reference system Liouville operator $iL_{\mathrm{ref}}^{(x)}$ up to time $t'$ starting from initial conditions $x'$, $\dot{x}'$ and $\Gamma_x'$ with an analogous meaning for $\dot{x}_{\mathrm{ref}}[x', \dot{x}', \Gamma_x'; t']$. Similarly, $y_{\mathrm{adb}}[y', \dot{y}', \Gamma_y', x'; t']$ indicates the exact adiabatic, *i.e.*, at fixed $x$, evolution of $y$ up to time $t'$ under the action of the operator in brackets in Eq. (11) starting from initial conditions $y'$, $\dot{y}'$, $\Gamma_y'$. Since $m_x \gg m_y$, the motion of the $y$ variable will follow instantaneously the motion of $x$ dynamically and sample its available configuration space over the time interval $\Delta t/2$ during which $x$ remains fixed. In Eq. (11), it can be seen that the force governing the evolution of $x$ and $\dot{x}$ is related to a time average of $F_x$ over the adiabatic $y$ trajectory.

Assuming that the $y$ motion is ergodic over this time interval, then the time average of $F_x$ appearing in Eq. (11) can be replaced by a configurational average over $y$ according to

$$\begin{aligned}
&\int_\tau^{\tau+\Delta t/2} dt F_x[y_{\mathrm{adb}}(y(\tau), \dot{y}(\tau), \Gamma_y(\tau), x; t), x] \\
&\quad = \frac{\Delta t}{2}\frac{\int dy F_x(x,y) e^{-\beta_y V(x,y)}}{\int dy\ e^{-\beta_y V(x,y)}} \\
&\quad = \frac{\Delta t}{2}\frac{\partial}{\partial x}\frac{1}{\beta_y}\ln Z_y(x; \beta_y)
\end{aligned} \tag{12}$$

where $\beta_y = 1/kT_y$ and

$$Z_y(x;\beta_y) = \int dy\ e^{-\beta V(x,y)} \tag{13}$$

is an effective configurational partition function of $x$. Note that the integration over $p_y$ trivially cancels out of Eq. (12). Therefore, in the adiabatic limit, an effective Hamiltonian for the $x$ degree of freedom can be constructed according to

$$H_x(p_x,x) = \frac{p_x^2}{2m_x} - \frac{1}{\beta_y}\ln\, Z_y(x;\beta_y) \equiv K_x - \frac{1}{\beta_y}\ln\, Z_y(x;\beta_y) \tag{14}$$

From Eq. (14) follow the statistical mechanical properties of the $x$ degree of freedom. The partition function for $x$ is given by

$$\begin{aligned} Q(\beta_x;\beta_y) &= \int dp_x dx \exp(-\beta_x H_x(p_x,x)) \\ &= \int dp_x \exp(-\beta_x K_x) \int dx \exp\left[-\beta_x\left(-\frac{1}{\beta_y}\ln\, Z_y(x;\beta_y)\right)\right] \\ &= \int dp_x K_x \int dx [Z_y(x;\beta_y)]^{\beta x/\beta_y} \end{aligned} \tag{15}$$

Therefore, the probability distribution function of $x$ becomes

$$P(x) = \frac{1}{Q(\beta_x,\beta_y)} \int dp_x \exp(-\beta_x K_x)[Z_y(x;\beta_y)]^{\beta_x/\beta_y} \tag{16}$$

so that $P(x) \propto [Z_y(x;\beta_y)]^{\beta_x/\beta_y}$. From Eq. (16), it follows that the free energy profile, $F(x)$, which, by definition is

$$F(x) = -\frac{1}{\beta_y}\ln\, Z_y(x;\beta_y) \tag{17}$$

can be calculated directly from the adiabatically generated probability distribution function, $P(x)$ by

$$F(x) = -\frac{1}{\beta_x}\ln\, P(x) + \text{const} \tag{18}$$

The equivalence between Eqs. (17) and (18) can be seen by direct substitution of $P(x)$ into Eq. (18). Thus, the true free energy profile can be obtained, up to an irrelevant constant, from the probability distribution function, $P(x)$, generated by the adiabatic dynamics with $m_x \gg m_y$ and

$T_x \gg T_y$ *via* Eq. (18). Note that Eq. (18) is of the same form as the standard free energy profile defined in Eq. (1) with $T$, the temperature of the ensemble, replaced by $T_x$ the temperature to which the $x$ degree of freedom has been heated. Despite the fact that $T_x$ appears in Eq. (18), the *correct* free energy profile at the temperature, $T_y$ of the bath/environment is obtained. Equation (18) constitutes the central result of this paper.

## 3. MODEL PROBLEMS AND RESULTS

In this section, the adiabatic free energy dynamics (AFED) method will be tested on two model problems. The first is a simple two-variable problem of the type described by Eq. (3) for which the free energy profile can be solved analytically. The second is a simple isomerization reaction in a Lennard-Jones solvent, which serves as a paradigm for solution phase chemical processes. In the second case the AFED method will also be compared to the "bluemoon ensemble" approach based on constrained molecular dynamics [5, 6].



### 3.1. One-dimensional Quartic Double Well Coupled to a Harmonic Oscillator

Consider a simple two-variable system described by the Hamiltonian in Eq. (3) with a potential of the form

$$V(x,y) = D_0(x^2 - a^2)^2 + \frac{1}{2}\kappa y^2 + \lambda xy \tag{19}$$

For this simple problem, $Z_y(x;\beta_y)$ can be calculated analytically, leading to a free energy profile in $x$ given by

$$F(x) = -\frac{1}{\beta_y}\ln Z_y(x;\beta_y) = D_0(x^2 - a^2)^2 - \frac{\lambda^2}{2\kappa}x^2 \tag{20}$$

Simulations of this model system were carried out using $D_0 = 5$, $a = 1$, $\kappa = 1$, $\lambda = 2.878$, $m_y = 1$ and $kT_y = 1$. With these parameters, the free energy profile in $x$ has two wells located at $x_\pm = \pm\sqrt{a^2 + \lambda^2/4D_0\kappa} = \pm 1.189$ and a barrier at $x = 0$. In addition, the free energy barrier is $F^\ddagger = F(x=0) - F(x_\pm) = D_0a^4 + \lambda^2a^2/2\kappa + \lambda^4/16D_0\kappa^2 = 10$.

In order to ensure efficient barrier crossing, a temperature $kT_x = 10$ was chosen for AFED simulations. With this choice of $kT_x$, the convergence of

the free energy profile with $m_x$ was tested by performing simulations with $m_x = 10$, 100, 300. For these mass choices, simulations of $10^9$ steps using a time step $0.25 \times 10^{-3}$ were performed. Canonical sampling is obtained using the recently introduced generalized Gaussian moment thermostat(GGMT) algorithm [13]. In general, the thermostatting method can have an influence on the efficiency of free energy methods. A detailed comparison between the GGMT approach and the more standard Nosé–Hoover chain algorithm will be given elsewhere [15].

Figure 1 shows the trajectory of $x$ as a function of time for $m_x = m_y$ and $T_x = T_y$ corresponding to an ordinary dynamics simulations compared to the AFED parameters ($m_x = 300$ and $T_x = 10$). The figure shows that without the adiabaticity conditions, barrier crossing is a rare event as would be expected in an ordinary dynamics calculation. In contrast, the AFED dynamics case shows frequent barrier crossing and, hence, efficient sampling of the configuration space available to $x$. Figure 2 shows the free energy profiles obtained from the AFED simulations for the different choices of $m_x$ together with the analytical result. The figure shows that when $m_x$ is too small, adiabaticity is not well maintained and the free energy profile is not well reproduced. It can be seen that for $m_x = 300$ the agreement between the AFED and analytical results is very good. In Section 4, a general protocol for determining the adiabaticity control parameters, $T_x$ and $m_x$ is discussed.



### 3.2. Isomerization Reaction in a Lennard-Jones Solvent

As a simple paradigm for solution phase chemical processes, we consider a simple isomerization reaction of a diatomic molecule solvated in a Lennard-Jones liquid. The intramolecular potential of the diatomic is characterized by a single coordinate $r = |\mathbf{r}_1 - \mathbf{r}_2|$, the distance between the two atoms comprising the diatomic. The intramolecular potential is, again, given by a quartic double well form:

$$V_{\text{intra}}(r) = D_0[(r - r_0)^2 - a^2]^2 \tag{21}$$

so that the total potential is given by

$$U(\mathbf{r}_1, \ldots, \mathbf{r}_N) = V_{\text{intra}}(|\mathbf{r}_1 - \mathbf{r}_2|) + V_{\text{LJ}}(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{22}$$

In this case, the potential parameters are $D_0 = 10^7\,\text{K/Å}^4$, $a = 0.22\,\text{Å}$, and $r_0 = 4.26\,\text{Å}$. With these parameters, the diatomic can exist in two stable conformations characterized by bond lengths of $r_\pm = r_0 \pm a$, which leads to $r_+ = 4.48\,\text{Å}$ and $r_- = 4.04\,\text{Å}$. The two conformations are separated by a

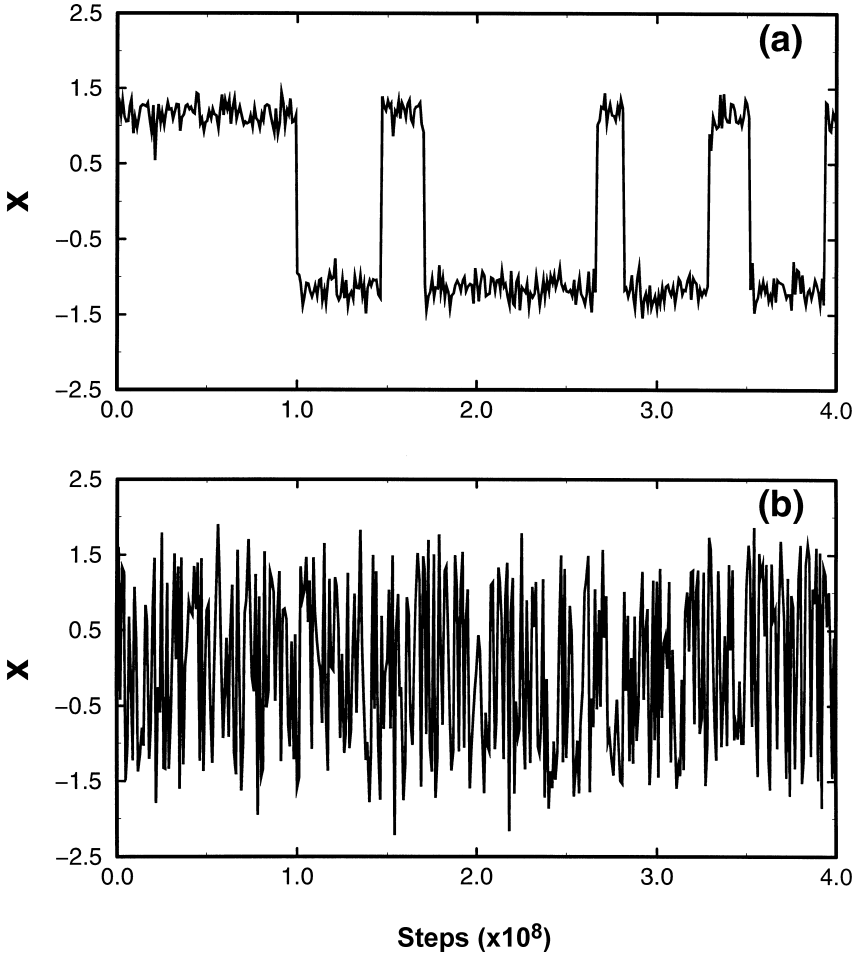




FIGURE 1 Trajectory of $x$ as a function of the number of steps for a double well coupled to a harmonic oscillator in Eq. (19). (a) Standard Molecular Dynamics; (b) Adiabatic (AFED) dynamics with $m_x = 300$ and $T_x = 10$.

46.5 kcal/mol potential barrier at $r = r_0$. The diatomic is solvated in a bath of 108 Lennard-Jones particles with $\varepsilon/k = 90\,\mathrm{K}$, $\sigma = 3.405\,\text{Å}$ and mass $m = 39.94$ amu. The masses of the two atoms in the diatomic are both the same as the mass of the Lennard-Jones particles. The density and temperature of the system are $\rho\sigma^3 = 0.844$ and $T = 300\,\mathrm{K}$.

In this example, the reaction coordinate is $r$, the diatomic bond length. The AFED method requires that the equations of motion be formulated in a coordinate system in which $r$ is an explicit coordinate and is given both a

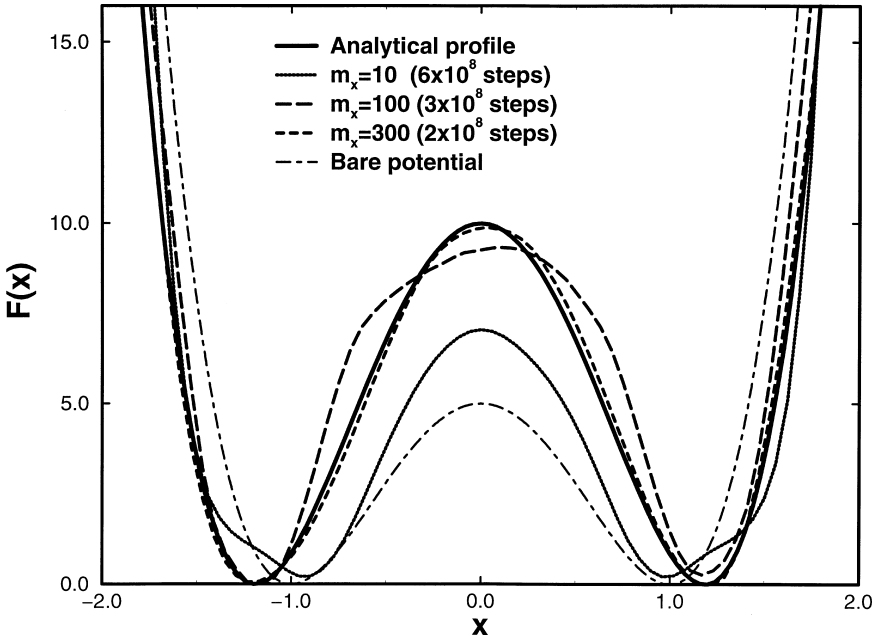


FIGURE 2 Adiabatic (AFED) Dynamics free energy profile for a double well coupled to a harmonic oscillator in Eq. (19). The figure shows the convergence of the free energy profile as a function of $m_x$ with $T_x = 10$. The solid line is the analytical free energy profile, the dotted line corresponds to $m_x = 10$, the long dashed line corresponds to $m_x = 100$, and the short dashed line corresponds to $m_x = 300$. $m_y = 1$ in all cases. For comparison, the bare potential is shown with the dot-dashed line.



large mass and temperature. To see how this is accomplished, consider the Lagrangian of the system:

$$L = \frac{1}{2}m(\dot{\mathbf{r}}_1^2 + \dot{\mathbf{r}}_2^2) + K_{\text{bath}} - V_{\text{intra}}(|\mathbf{r}_1 - \mathbf{r}_2|) - V_{\text{LJ}}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_N) \tag{23}$$

where $K_{\text{bath}}$ is the kinetic energy of the Lennard-Jones bath particles. First, we transform to center of mass and relative coordinates for the diatomic, $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$ and $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, yielding:

$$L = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2 + K_{\text{bath}} - V_{\text{intra}}(|\mathbf{r}|) - V_{\text{LJ}}\left(\mathbf{R} + \frac{1}{2}\mathbf{r}, \mathbf{R} - \frac{1}{2}\mathbf{r}, \mathbf{r}_3, \ldots, \mathbf{r}_N\right) \tag{24}$$

where $M = 2m$ and $\mu = m/2$ are the total and reduced masses of the diatomic, respectively. Finally, we transform to spherical polar coordinates $(r, \theta, \phi)$ in

the relative coordinate giving:

$$L = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu r^2\dot{\mathbf{u}}^2 + K_{\text{bath}} - V_{\text{intra}}(r) - V_{\text{LJ}}\left(\mathbf{R} + \frac{1}{2}r\mathbf{u}, \mathbf{R} - \frac{1}{2}r\mathbf{u}, \mathbf{r}_3, \ldots, \mathbf{r}_N\right) \quad (25)$$

where the unit vector $\mathrm{u} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the bond orientation vector. The adiabaticity condition is imposed by choosing a temperature $T_r$ associated with the radial degree of freedom only and by redefining the Lagrangian in Eq. (25) according to

$$L = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\tilde{\mu}\dot{r}^2 + \frac{1}{2}\mu r^2\dot{\mathbf{u}}^2 + K_{\text{bath}} - V_{\text{intra}}(r) - V_{\text{LJ}}\left(\mathbf{R} + \frac{1}{2}r\mathbf{u}, \mathbf{R} - \frac{1}{2}r\mathbf{u}, \mathbf{r}_3, \ldots, \mathbf{r}_N\right) \quad (26)$$

where $\tilde{\mu} \gg \mu$ is a mass associated with the $r$ degree of freedom only.

It is important to note that, although it is necessary to work with the coordinate $r$ explicitly, it is *not* necessary to work directly with the angular degrees of freedom $\theta$ and $\phi$. Rather, we work with the Cartesian components $(u_x, u_y, u_z)$ the vector $\mathbf{u}$ directly so that the kinetic energy term involving $\mathbf{u}$ becomes:

$$\frac{1}{2}\mu r^2\dot{\mathbf{u}}^2 = \frac{1}{2}\mu r^2(\dot{u}_x^2 + \dot{u}_y^2 + \dot{u}_z^2) \quad (27)$$

When this is done, the equations of motion generated by Eq. (26) will be in terms of $r$ and the three Cartesian components of $\mathbf{u}$. In order to ensure that $\mathbf{u}$ remains a unit vector as the system evolves in time, it is necessary to add a simple constraint to the equations of motion that $u_x^2 + u_y^2 + u_z^2 = 1$. Details of how the equations of motion are integrated and how the constraint is implemented are given in Appendix B.

AFED simulations were carried out using temperatures $T_r = 200T = 60000\,\text{K}$ and $T_r = 30T = 9000\,\text{K}$. At $T_r = 200T$, masses of $m_r \equiv \tilde{\mu} = 3000m \approx 120 \times 10^3\,\text{amu}$ and $m_r = \tilde{\mu} = 6000m \approx 240 \times 10^3\,\text{amu}$ were used with a time step of $\Delta t = 5\,\text{fs}$. At $T_r = 30T$, masses of $m_r \equiv \tilde{\mu} = 11984\,\text{amu}$ and 35953 amu were used. AFED simulation lengths of various lengths were performed in order to test convergence. In addition, bluemoon ensemble simulations were carried out using 11 evenly spaced fixed values of $r$ between 4.0 Å and 4.50 Å. (This was found to be the minimum number of points needed to generate an accurate free energy profile.) Each bluemoon simulation consisted of $2 \times 10^4$ steps with a time step of $\Delta t = 10\,\text{fs}$ for a total run time of 2.2 ns. Although this is somewhat longer than is actually

needed to generate the free energy profile *via* the bluemoon ensemble method, such a long run ensures a highly converged free energy profile against which the AFED method can be benchmarked. In Figure 3, the free

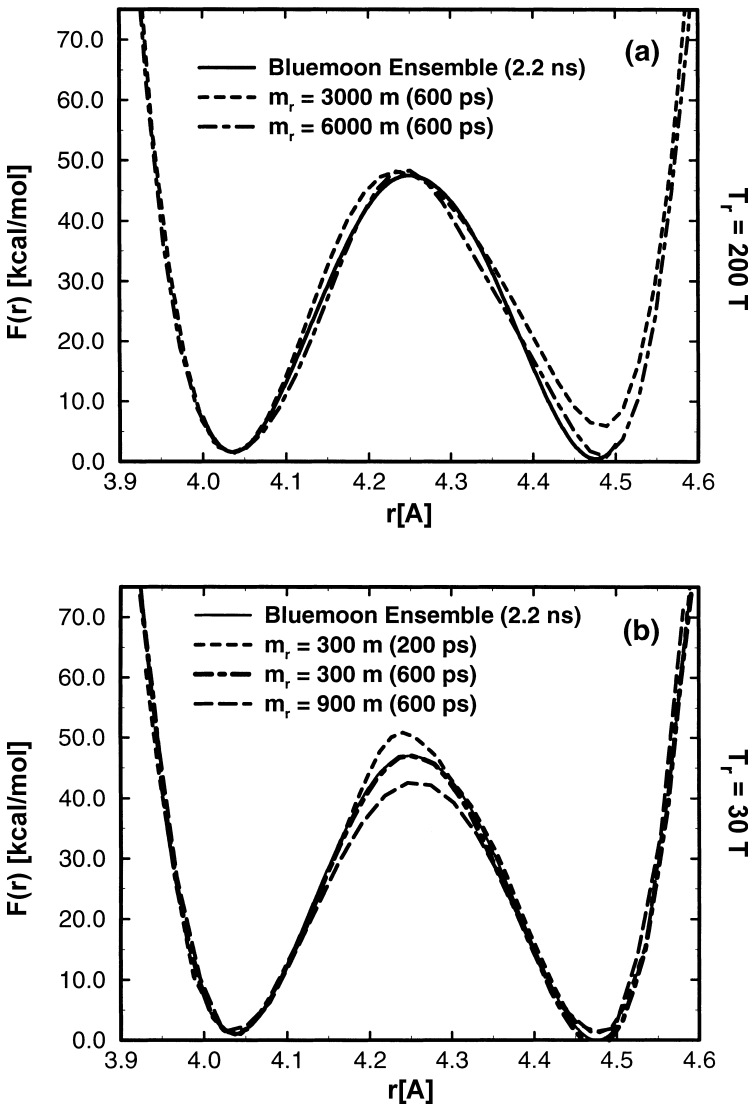


FIGURE 3 Adiabatic Dynamics free energy profile for an isomerization reaction of a diatomic in a Lennard-Jones solvent described in Section 3.2. The figure shows the convergence as a function of the mass $\tilde{\mu}$ (*cf.* Eq. (26)) with $T_r = 200T = 60000\,\mathrm{K}$ (a) and for $T_r = 30T = 9000\,\mathrm{K}$ (b). In both panels, the solid line is the free energy profile obtained employing the "bluemoon ensemble" method [5, 6]. In all cases $m = 39.94\,\mathrm{amu}$ and the temperature and density are $T = 300\,\mathrm{K}$ and $\rho\sigma^3 = 0.84$, respectively.

energy profiles obtained from the AFED calculations for the various mass and temperature choices and that obtained from the bluemoon calculation are shown. It can be seen that for $T_r = 200\,T$ and $m_r \equiv \tilde{\mu} = 240 \times 10^3$ amu, the agreement between the two methods is good, and for $T_r = 30\,T$ and $m_r \equiv \tilde{\mu} = 300\,m = 11984$ amu, the agreement between the two methods is also good. In the latter case, a 600 ps run yields the best agreement for $m_r = 300\,m$, however, even a 200 ps run reproduces the free energy barrier to within about 5%. For $m_r = 900\,m$, a 600 ps run is insufficiently long to converge the free energy profile, as can be see in Figure 3(b), however, if the trajectory is allowed to run for 2 ns, it is found that the correct free energy profile is obtained.

In order to compare the efficiency of the AFED and bluemoon methods, the following analysis was carried out. First, note in Figure 3, that the AFED method actually generates a larger range of $r$ than the bluemoon method. Of course, the bluemoon method can be made to sample the same range of $r$ by carrying out more simulations. In this case, 22 evenly spaced points would be required, hence, 22 simulations. Next, an AFED simulation using 22 bins for the histogram of $P(r)$ was carried out, and the error bar on the force $F_r$ on $r$ in one of the bins computed and compared to the error bar obtained from the bluemoon ensemble method for the same value of $r$. It was found that, in order to obtain an error bar of size $8 \times 10^{-5}$ au, an AFED simulation of 880 ps was needed compared to a bluemoon calculation of 80 ps. However, since one AFED simulation generates the full free energy profile, while 22 individual bluemoon simulation are needed, the total simulation time for the bluemoon method would then be $22 \times 80\,\text{ps} = 1.76$ ns. Thus, comparing the total simulation times, shows that the AFED method is nearly twice as efficient as the bluemoon method. In practice, one might be content with a restricted range for the reaction coordinate in the bluemoon method. In the present example, the above analysis shows that, in this case, the two method are, then, of equal efficiency, however, the AFED method gives a more complete picture of the free energy profile.



## 4. DISCUSSION: CHOOSING THE ADIABATICITY PARAMETERS

In this section, a protocol for choosing the parameters for the reaction coordinate temperature and mass in order to ensure that adiabaticity is maintained and that phase space is properly sampled will be discussed. The

temperature of the reaction coordinate needs to be close to the barrier height of the bare potential divided by Boltzmann's constant, however, as Figure 3 shows, it is possible for it to be greater or slightly less than this value. In cases where free energy barriers are expected to be higher than the bare potential barrier, a larger temperature will likely be needed. In general, higher temperatures work as well but require higher masses. Typically, a short run is sufficient to determine if the reaction coordinate is undergoing frequent barrier crossing events. Note that the higher the temperature of the reaction coordinate, the larger will be the required mass separation between the reaction coordinate and other modes in order to maintain adiabaticity. The mass of the reaction coordinate, which controls the adiabaticity, needs to be chosen such that the characteristic frequency of the reaction coordinate motion is small compared to that of the remaining degrees of freedom. In general, it is possible to provide estimates of these frequencies for the physical system under consideration. Generally, it is found that a frequency scale separation of approximately 2–4 is sufficient to maintain satisfactory adiabatic separation. As is clear from the parameters used in the two examples of Section 3, such frequency scale separations lead to the most accurate results for the shortest simulation runs. Note that once a mass scale separation for the reaction coordinate is chosen, the time scale for the evolution of the dynamical thermostats needs to be adjusted accordingly (see Refs. [12, 13]).



## 5. CONCLUSION

Given the importance of free energy calculations in the study of rare events, it is important to develop novel and efficient methods for computing free energy profiles along reaction paths. In this paper, a new approach to the determination of free energy profiles has been introduced. The method is based on the creation of an adiabatic separation between the reaction coordinate and the remaining degrees of freedom within a molecular dynamics simulation. In addition, the reaction coordinate is maintained at a high temperature relative to the remaining degrees of freedom. In this way, the full configuration space corresponding to the rare event is sampled and the free energy profile is rigorously generated directly from the probability distribution of the reaction coordinate. Thus, the new method requires no postprocessing of the data and leads to a scheme that is more efficient than methods based on constraining (or restraining) the reactive degree of freedom.

The new AFED method can be employed in any situation to which other free energy methods can be applied. In particular, in complex biomolecular applications, it offers several advantages in addition to increased efficiency. First, it requires no "by-hand" adjustments of the reaction coordinate. Such adjustments, usually needed in the bluemoon ensemble and umbrella sampling methods, can often be difficult to perform for complex reaction coordinates and/or reaction coordinates that are strongly coupled to other degrees of freedom. A good example is that of one or more internal dihedral angle(s). Second, when a definite reaction coordinate is not available, it is expected that the AFED method will allow the free energy profile of a subspace of generalized coordinates to be obtained with greater efficiency than other free energy methods. Testing this hypothesis will be part of future work to be carried out in this area.

Other future work will include adapting the method for other types of free energy calculations (*e.g.*, solvation free energies) and combining the AFED method with path integral molecular dynamics [16] for calculation of quantum free energy profiles. The latter will allow, for example, exploration of equilibrium and kinetic isotope effects.



***Acknowledgements***

This work has been funded by ACS-PRF #33256-G6, Research Corporation #RI0218, and NSF CHE-98-75824.

## APPENDIX A

In this appendix, the problem of treating an $N$-particle system in a set of generalized coordinate that explicitly contains the reaction coordinate is considered. From a dynamical point of view, if one begins with the Lagrangian in Cartesian coordinates, $\mathbf{r}_1, \ldots, \mathbf{r}_N$

$$L = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - V(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{A.1}$$

then, under a transformation to generalized coordinates $q_1, \ldots, q_{3N}$

$$\begin{aligned} q_1 &= q_1(\mathbf{r}_1, \ldots, \mathbf{r}_N) \\ &\cdots \\ q_{3N} &= q_{3N}(\mathbf{r}_1, \ldots, \mathbf{r}_N) \end{aligned} \tag{A.2}$$

the Lagrangian becomes

$$L = \sum_{\alpha,\beta=1}^{3N} \frac{1}{2} G_{\alpha\beta}(q)\dot{q}_\alpha \dot{q}_\beta - V(\mathbf{r}_1(q), \ldots, \mathbf{r}_N(q)) \tag{A.3}$$

where $q \equiv \{q_1, \ldots, q_{3N}\}$ denotes the full set of generalized coordinates, and $G_{\alpha\beta}(q)$ is given by

$$G_{\alpha\beta}(q) = \sum_{i=1}^{3N} m_i \left( \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \right) \cdot \left( \frac{\partial \mathbf{r}_i}{\partial q_\beta} \right) \tag{A.4}$$

is the mass-metric tensor. By performing the Legendre transform, the Hamiltonian can be shown to be

$$H = \frac{1}{2} \sum_{\alpha,\beta=1}^{3N} K_{\alpha\beta}(q) p_\alpha p_\beta + V(\mathbf{r}_1(q), \ldots, \mathbf{r}_N(q)) \tag{A.5}$$

where

$$K_{\alpha\beta}(q) = \sum_{i=1}^{N} \frac{1}{m_i} \left( \frac{\partial q_\alpha}{\partial \mathbf{r}_i} \right) \cdot \left( \frac{\partial q_\beta}{\partial \mathbf{r}_i} \right) \tag{A.6}$$

is the inverse of $G_{\alpha\beta}$. The introduction of coordinate-dependent mass-metric factors into the Lagrangian/Hamiltonian of a system *via* a transformation to generalized coordinates leads to a considerable increase in complexity in the description of the true dynamics of the system.

Although the AFED method requires that generalized coordinates be used, the dynamics only needs to generate the correct configurational averages, and, hence, it is not necessary to work with correct conjugate momenta. This fact leads to a large simplification in the analysis and implementation of the method. In order to show this, consider the canonical partition function of the system. In Cartesian coordinates, this is given by

$$Q(N,V,T) = \frac{1}{N!h^{3N}} \int d^N\mathbf{p} d^N\mathbf{r} \exp\left[ -\beta \left( \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_1, \ldots, \mathbf{r}_N) \right) \right] \tag{A.7}$$

If the coordinate integration only is transformed according to Eq. (A.2), leaving the momentum integrals unchanged, the partition function picks up

a Jacobian factor according to

$$Q(N,V,T) = \frac{1}{N!h^{3N}} \int d^N\mathbf{p}\; d^{3N}q J(q) \exp\left[ -\beta\left( \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_1(q),\ldots,\mathbf{r}_N(q)) \right)\right]$$

$$= \frac{1}{N!h^{3N}} \int d^N\mathbf{p}\; d^{3N}q \times \exp\left[ -\beta\left( \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_1(q),\ldots,\mathbf{r}_N(q)) - \frac{1}{\beta}\ln J(q) \right)\right] \tag{A.8}$$

Defining an effective potential according to

$$\tilde{V}(q_1,\ldots,q_{3N};\beta) = V(\mathbf{r}_1(q),\ldots,\mathbf{r}_N(q)) - \frac{1}{\beta}\ln\, J(q_1,\ldots,q_{3N}) \tag{A.9}$$

the partition function becomes

$$Q(N,V,T) = \frac{1}{N!h^{3N}} \int d^N\mathbf{p}\; d^{3N}q \exp\left[ -\beta\left( \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \tilde{V}(q_1,\ldots,q_{3N}) \right)\right] \tag{A.10}$$

It is, thus, clear that the phase space distribution function in Eq. (A.10) can be generated *via* molecular dynamics with a Hamiltonian of the form

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \tilde{V}(q_1,\ldots,q_{3N}) \tag{A.11}$$

where the $3N$ Cartesian momenta are treated as "conjugate" to the $q$s even though they are not the true conjugate momenta. Similarly, if $q_1$ is a reaction coordinate of interest, then the AFED method can be implemented by introducing a temperature $T_1$ for this variable and writing the Hamiltonian in the form

$$H = \frac{p_1^2}{2\tilde{m}_1} + \sum_{\alpha=2}^{3N} \frac{p_\alpha^2}{2m_\alpha} + \tilde{V}(q_1,\ldots,q_{3N}) \tag{A.12}$$

where $m_1 \gg m_\alpha$, $\alpha = 2,\ldots,3N$, in order to ensure an adiabatic separation. Under these conditions, the analysis of Section 2 can be applied to a general $N$-particle system in generalized coordinates with no modification except for

a simple replacement of the $N$-particle potential $V$ with the modified potential $\tilde{V}$. It should be noted that, under certain conditions, it may be straightforward to work in terms of a full canonical set of generalized coordinates, such as in the example of Section 3.2.

## APPENDIX B

In this appendix, the implementation of the AFED method for the isomerization reaction of Section 3.2 is discussed. The procedure presented here is generally applicable to any problem for which the reaction coordinate is a distance. More general procedures for other types of reaction coordinates have been discussed in Appendix A. As seen in Section 3.2 the Lagrangian of the system can be rewritten as:

$$L = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\tilde{\mu}\dot{r}^2 + \frac{1}{2}\mu r^2\dot{\mathbf{u}}^2 + K_{\text{bath}} - V_{\text{intra}}(r) - V_{\text{LJ}}\left(\mathbf{R} + \frac{1}{2}r\mathbf{u}, \mathbf{R} - \frac{1}{2}r\mathbf{u}, \mathbf{r}_3, \ldots, \mathbf{r}_N\right) \tag{B.1}$$

where $r$, the bond length, is the reaction coordinate, $\mathbf{u}$ is the unit vector along the bond, $K_{\text{bath}}$ is the kinetic energy of the bath, and $\tilde{\mu} \gg \mu$ is a mass associated with the $r$ degree of freedom only. In addition, $F_r = -V'_{\text{intra}}(r) - \partial V_{\text{LJ}}/\partial r$ and $\mathbf{F}_{\mathbf{u}} = -\partial V_{\text{LJ}}/\partial \mathbf{u}$. From Eq. (B.2), the equations of motion for $r$ and $\mathbf{u}$ follow directly:

$$\dot{r} = v_r \tag{B.2}$$

$$\dot{\mathbf{u}} = \mathbf{v}_{\mathbf{u}} \tag{B.3}$$

$$\dot{v}_r = \frac{\mu}{\tilde{\mu}} r\mathbf{v}_{\mathbf{u}}^2 + \frac{1}{\tilde{\mu}}F_r \tag{B.4}$$

$$\dot{\mathbf{v}}_{\mathbf{u}} = -\frac{2}{r}v_r\mathbf{v}_{\mathbf{u}} + \frac{1}{\mu r^2}\mathbf{F}_{\mathbf{u}} \tag{B.5}$$

The equations of motion, Eq. (B.5), are slightly more complicated than the usual equations of motion encountered in ordinary molecular dynamics, and, therefore, the problem of constructing a reversible, symplectic integrator for them based on the Liouville operator formalism is discussed below.

The Liouville operator corresponding to Eq. (B.5) is

$$\begin{aligned} iL &= iL_1 + iL_2 + iL_3 \\ iL_1 &= v_r \frac{\partial}{\partial r} + \mathbf{v_u} \cdot \frac{\partial}{\partial \mathbf{u}} \\ iL_2 &= \left( \frac{F_r}{\tilde{\mu}} + \frac{\mu}{\tilde{\mu}} r \mathbf{v}_\mathbf{u}^2 \right) \frac{\partial}{\partial v_r} \\ iL_3 &= \left( -\frac{2}{r} v_r \mathbf{v_u} + \frac{1}{\mu r^2} \mathbf{F_u} \right) \cdot \frac{\partial}{\partial \mathbf{v_u}} \end{aligned} \tag{B.6}$$

In order to generate a reversible integrator for Eq. (B.5), the following factorization of the classical propagator, $\exp(iL\Delta t)$ for a time step, $\Delta t$ is constructed:

$$\begin{aligned} \exp(iL\Delta t) \approx & \exp\left( iL_2 \frac{\Delta t}{4} \right) \times \exp\left( iL_3 \frac{\Delta t}{2} \right) \times \exp\left( iL_2 \frac{\Delta t}{4} \right) \\ & \times \exp(iL_1 \Delta t) \\ & \times \exp\left( iL_2 \frac{\Delta t}{4} \right) \times \exp\left( iL_3 \frac{\Delta t}{2} \right) \times \exp\left( iL_2 \frac{\Delta t}{4} \right) \end{aligned} \tag{B.7}$$

Application of this operator to the phase space vector gives the time evolution of the variables, $r$, $v_r$, $\mathbf{u}$, $\mathbf{v_u}$. The operators that act on the variables $r$, $v_r$, $\mathbf{u}$ yield a simple time translation, while the action of $\exp(iL_3 \Delta t/2)$ gives the following evolution of $\mathbf{v_u}$:

$$\mathbf{v_u}(\Delta t/2) = \mathbf{v_u}(0) \exp\left( -\frac{1}{r} v_r \Delta t \right) + \frac{1}{2\mu r v_r} \left( 1 - \exp\left( -\frac{1}{r} v_r \Delta t \right) \right) \mathbf{F_u}(0) \tag{B.8}$$

where $v_r$ has the value it has obtained when the operator, $\exp(iL_3 \Delta t/2)$ is applied. The potentially singular term, $(1 - \exp(-(2/r)v_r t))/v_r$, can be expanded in a Maclaurin series when $v_r$ is small, since its limit for $v_r \to 0$ is finite.

In order to complete the scheme, it is necessary to add an additional force that constrains $\mathbf{u}$ to be a unit vector. This is accomplished by modifying the equation of motion for $\mathbf{u}$ according to read

$$\dot{\mathbf{v}}_\mathbf{u} = -\frac{2}{r} v_r \mathbf{v_u} + \frac{1}{\mu r^2} (\mathbf{F_u} - \lambda \mathbf{u}) \tag{B.9}$$

In Eq. (B.9), a constraint force $-\lambda \mathbf{u}$ has been added to $\mathbf{F_u}$, where $\lambda$ is a Lagrange multiplier that ensures the condition $\mathbf{u}(t) \cdot \mathbf{u}(t) = 1$ is maintained at all time $t$. The multiplier is calculated using the standard procedure. The

vector $\mathbf{u}$ is updated in the absence of the constraint force using Eq. (B.8) to produce the unconstrained evolution $\mathbf{u}^u(\Delta t)$. The vector $\mathbf{u}(\Delta t)$ is then constructed according to

$$\mathbf{u}(\Delta t) = \mathbf{u}^u(\Delta t) - \lambda \frac{\Delta t^2}{2\mu r^2}\mathbf{u}(0) \tag{B.10}$$

Defining $\lambda' = \lambda \Delta t^2 / 2\mu r^2$, the condition $|\mathbf{u}(\Delta t)|^2 = 1$ is then imposed, which leads to the following expression for $\lambda'$:

$$\lambda' = \mathbf{u}(0) \cdot \mathbf{u}^u(\Delta t) \sqrt{|\mathbf{u}(0) \cdot \mathbf{u}^u(\Delta t)|^2 - (|\mathbf{u}^u(\Delta t)|^2 - 1)} \tag{B.11}$$

Once the multiplier has been determined, $\mathbf{u}(\Delta t)$ is obtained from Eq. (B.10), and the velocity $\mathbf{v}_\mathbf{u}(\Delta t/2)$ at the half step obtained from

$$\mathbf{v}_\mathbf{u}(\Delta t/2) = \mathbf{u}^u(\Delta t/2) - \frac{\lambda'}{\Delta t}\mathbf{u}(0) \tag{B.12}$$

The forces are then recalculated at the new positions, and $\mathbf{v}_\mathbf{u}$ is updated with the appropriate updated unconstrained force according to Eq. (B.8). Finally, the constraint force is applied, and the resulting update for $\mathbf{v}_\mathbf{u}$, leading to the fully updated velocity is

$$\mathbf{v}_\mathbf{u}(\Delta t) = \mathbf{u}_\mathbf{u}^u(\Delta t) - (\mathbf{u}(\Delta t) \cdot \mathbf{v}_\mathbf{u}^u(\Delta t))\mathbf{u}(\Delta t) \tag{B.13}$$

where $\mathbf{v}_\mathbf{u}^u(\Delta t)$ is the velocity just before the constraint force at $\Delta t$ is applied. Additionally, it is necessary to apply thermostats to ensure canonical sampling. As discussed in Section 2, a thermostat is applied to the reaction coordinate $r$, maintaining its temperature at an elevated value $T_r$. Because of the constraint on $\mathbf{u}$, it is necessary to thermostat the three components of $\mathbf{u}$ together and separately from the remaining bath degrees of freedom [17]. Apart from this, the application of the thermostat operators follows the procedure described in detail in Refs. [13] and [17]. Finally, it should be noted that the formulation of the integrator *via* the Liouville operator allows the AFED method to be easily combined with multiple time scale integration (RESPA) [14] techniques for the bath.



## *References*

[1] Radmer, R. J. and Kollman, P. A. (1997). Free energy calculations: A theoretical and empirical comparison of numerical errors and a new method for qualitative estimates of free energy changes. *J. Comp. Chem.*, **18**, 902.

[2] Torrie, G. M. and Valleau, J. P. (1974). Monte-Carlo free energy estimates using non-Boltzmann sampling. Application to the subcritical Lennard-Jones fluid. *Chem. Phys. Lett.*, **28**, 578.

[3] Torrie, G. M. and Valleau, J. P. (1977). Nonphysical sampling distributions in Monte Carlo free energy estimation: Umbrella sampling. *J. Comp. Phys.*, **23**, 187.
[4] Valleau, J. P. and Torrie, G. M. (1977). Plenum Press, New York.
[5] Carter, E. A., Ciccotti, G., Hynes, J. T. and Kapral, R. (1989). Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.*, **156**, 472.
[6] Sprik, M. and Ciccotti, G. (1998). Free energy from constrained molecular dynamics. *J. Chem. Phys.*, **109**, 7737.
[7] Kumar, S., Swendsen, R. H., Kollman, P. A. and Rosenberg, J. M. (1992). The weighted histogram analysis method for free energy calculations on biomolecules. *J. Comp. Chem.*, **13**, 1011.
[8] Bruccoleri, R. E. and Karplus, M. (1990). Conformational sampling using high-temperature molecular dynamics. *Biopolymers*, **29**, 3975.
[9] Samuelson, S. and Martyna, G. J. (1997). Solvent, force field, temperature and quantum effects on the folding free energy surface of blocked alanine tripeptide. *J. Chim. Phys.*, **94**, 1503.
[10] Cao, J. and Martyna, G. J. (1996). Adiabatic path integral molecular dynamics. *J. Chem. Phys.*, **104**, 2028.
[11] Marx, D., Tuckerman, M. E. and Martyna, G. J. (1999). Quantum dynamics *via* adiabatic *ab initio* centroid molecular dynamics. *Comp. Phys. Comm.*, **118**, 166.
[12] Martyna, G. J., Tuckerman, M. E. and Klein, M. L. (1992). Nosé-Hoover chains: The canonical ensemble *via* continuous dynamics. *J. Chem. Phys.*, **97**, 2635.
[13] Liu, Y. and Tuckerman, M. E. (2000). Generalized Gaussian moment thermostatting: A new continuous dynamical approach to the canonical ensemble. *J. Chem. Phys.*, **112**, 1685.
[14] Tuckerman, M. E., Martyna, G. J. and Berne, B. J. (1992). Reversible multiple time scale molecular dynamics. *J. Chem. Phys.*, **97**, 1990.
[15] Rosso, L. and Tuckerman, M. E., Free energy profiles *via* adiabatic molecular dynamics. *J. Chem. Phys.* (To be published).
[16] Tuckerman, M. E., Martyna, G. J., Klein, M. L. and Berne, B. J. (1993). Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals. *J. Chem. Phys.*, **99**, 2796.
[17] Martyna, G. J., Tuckerman, M. E., Tobias, D. J. and Klein, M. L. (1996). Reversible integrators for extended system molecular dynamics. *Mol. Phys.*, **87**, 1117.