Torsion Dynamics of Molecular Systems

Sebastian Reich

Konrad-Zuse Zentrum Berlin, Heilbronner Str. 10, D-10711 Berlin

Abstract

Based on the concept of free energy, we derive a Hamiltonian formulation for molecular dynamics in torsion space. The appropriate reaction coordinates for the free energy calculations are defined in terms of soft constraints as introduced by B.R. Brooks, J. Zhou, and S. Reich in the context of molecular dynamics. We consider a few simplifications that allow one to calculate the free energy analytically and to write the corresponding equations of motion as a constraint Hamiltonian system that can conveniently be discretized by the well-known SHAKE algorithm. The additional computational costs, compared to using the orginal force field and constraining bond-lengths and bond-angles to their equilibrium value (hard constraints), amount, in general, to less than a complete force evaluation. We show for a single butane molecule that our Hamiltonian formulation yields the correct Boltzmann distribution in the torsion angle while the original Hamiltonian together with hard constraints on the bond-lengths and bond-angles results in a much reduced transition rate between the *trans* and *cis* configuration.

1. INTRODUCTION

For classical molecular dynamics (MD), atomic trajectories obey the Hamiltonian equation of motion

$$
\begin{aligned}
\frac{d}{dt}q &= M^{-1}p \\
\frac{d}{dt}p &= -\nabla V(q)
\end{aligned}
\tag{1}
$$

where q is the vector containing all positions (in cartesian coordinates), p is the vector containing all conjugate momenta, M is a diagonal matrix of atomic masses replicated thrice, and $V(q)$ is the (empirical) potential energy function.[1–3] Standard numerical schemes for simulating the dynamical behavior of molecules are based on discrete time-stepping. Such numerical simulations are complicated by the presence of multiple time-scales.[2–4] Standard integrators, such as Verlet,[5] have to use time-steps which are small compared to the fastest time-scales. In most cases, those time-scales come from bonded interactions. Often the interesting dynamical phenomena of a molecule happen, however, on much slower time-scales and are primarily related to motions in the dihedral angles.[3] Thus it seems reasonable to average over the fastest degrees of motion and then to solve the reduced equations numerically. This allows one to use larger time-steps and the computation of the long-term dynamics of molecules could become feasible. Several methods for the removal of the bonded interactions have been suggested.[6–10] Typical computational methods use (hard) constraints that freeze the bond-lengths and/or the bond-angles to their equilibrium value.[11,12] However, when applied to the bond-angle bending, the resulting molecule becomes too rigid and transition rates are no longer reproduced correctly.[12] For that reason soft constraints were introduced[4,10,13] which maintain the flexibility of a molecule in terms of its bonds and bond-angles.

A completely different approach to long-time integration is to discretize the equations of motion (1) by an implicit method and to use a step-size that is large with respect to the bond vibrations.[14] However, it is not very well understood how poor numerical resolution of the bond vibrations effects the overall simulation results.

In this paper we derive the reduced equations of motion by first calculating the free energy in terms of appropriately chosen reaction coordinates. Based on these analytical considerations, we propose a constraint formulation that can be discretized by the standard SHAKE algorithm algorithm.[11] The additional costs for solving the modified constraint equations are, in general, less than a complete force evaluation. By means of a simple example we demonstrate that the modifications suggested in this paper seem to overcome the transition-rate-problem of standard constraint methods; i.e., reproduce transition rates in the torsion angles correctly.

2. THE EQUATIONS OF MOTION

Let us rewrite the equations of motions (1) as

$$\begin{aligned} \frac{d}{dt}q &= M^{-1}p \\ \frac{d}{dt}p &= -\nabla\tilde{U}(q) - \tilde{G}(q)^T K \tilde{g}(q) \end{aligned} \tag{2}$$

where $\tilde{G} = \tilde{g}_q$ and $\tilde{g}$ is the collection of functions $\tilde{g}_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$, with corresponding force constants K_{ii}, i.e.

$$\frac{1}{2}\tilde{g}(q)^T K \tilde{g}(q) = \frac{1}{2}\sum_i K_{ii}(\tilde{g}_i(q))^2 , \tag{3}$$

and K the m-dimensional diagonal matrix with entries K_{ii}. The potential (3) stands for covalent bond stretching, i.e. $\tilde{g}_i(q) = r - r_0$, bond-angle bending, i.e. $\tilde{g}_i(q) = \phi - \phi_0$, and improper dihedral angles, i.e. $\tilde{g}_i(q) = \psi - \psi_0$. The potential $\tilde{U}(q)$ contains the proper torsion potentials, the Lennard-Jones interactions, and the electrostatic interactions.

The potential (3) represents the fastest degrees of motion of a macromolecule. To remove those degrees of freedom, we have to calculate the free energy of our system in terms of properly chosen reaction coordinates. Typically the reaction coordinates are defined by setting the bond-lengths and bond angles to their equilibrium values; i.e., one assumes

$$\tilde{g}(q) = 0 . \tag{4}$$

However, this leads to non-realistic simulation results when (4) is applied to (1) as a constraint through SHAKE.[12] Here we suggest to use soft constraints instead.[4,10,13] These are defined by requiring that the gradient of the total potential energy

$$V(q) = \tilde{U}(q) + \frac{1}{2}\tilde{g}(q)^T K \tilde{g}(q)$$

with respect to the bond-lengths, bond-angles and improper dihedral angles vanishes; i.e.

$$0 = \nabla_{\tilde{q}_1} V(q)$$

with $\tilde{q}_1 := \tilde{g}(q)$. Premultiplying the resulting expression by the matrix K^{-1}, this leads to

$$0 = \tilde{g}(q) + K^{-1}[\tilde{G}(q)M^{-1}\tilde{G}(q)^T]^{-1}\tilde{G}(q)M^{-1}\nabla\tilde{U}(q) \tag{5}$$

where we have assumed that $\tilde{G}(q)M^{-1}\tilde{G}(q)^T$ is invertible. Then we define the new function $g : \boldsymbol{R}^n \to \boldsymbol{R}$ by

$$g(q) := \tilde{g}(q) + K^{-1}[\tilde{G}(q)M^{-1}\tilde{G}(q)^T]^{-1}\tilde{G}(q)M^{-1}\nabla\tilde{U}(q) \tag{6}$$

and the reduced dynamics of (1) will now be defined by the free energy of (1) on the constraint manifold

$$\mathcal{M} = \{(q,p) \in \boldsymbol{R}^{2n} : \, g(q) = 0, \, G(q)M^{-1}p = 0\,\}\,. \tag{7}$$

The manifold can be parameterized by the unconstrained dihedral angles, the external degrees of freedom, and their corresponding conjugate momenta. For simplicity, we refer to the reduced dynamics on $\mathcal{M}$ as the torsion dynamics of (1). The corresponding free energy will be derived in Section 5.

3. EQUATIONS OF MOTION IN LOCAL COORDINATES

Let us rewrite (1) for theoretical purposes as

$$\begin{aligned} \frac{d}{dt}q &= M^{-1}p \\ \frac{d}{dt}p &= -\nabla U(q) - G(q)^T K g(q) \end{aligned} \tag{8}$$

with

$$U(q) := V(q) - \frac{g(q)^T K g(q)}{2}\,. \tag{9}$$

Next we reformulate (8) in local coordinates (q_1, q_2) defined by

$$\begin{aligned} q_1 &= g(q) \\ q_2 &= b(q) \end{aligned} \tag{10}$$

where $b(q)$ is a vector valued function such that $B(q)M^{-1}G(q)^T = 0$, $B(q) = b_q(q)$, and the composed matrix $[G(q)^T \, B(q)^T]$ is invertible and well conditioned. The existence of such a coordinate system follows, at least locally, from the Frobenius Theorem.[15] The corresponding conjugate momenta are given by

$$[G(q)^T \, B(q)^T] \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = p \tag{11}$$

which results in the Hamiltonian

$$H(q,p) = \frac{p_1^T G M^{-1} G^T p_1}{2} + \frac{p_2^T B M^{-1} B^T p_2}{2} + U + \frac{q_1^T K q_1}{2}. \tag{12}$$

The equations of motion are now given by

$$\begin{aligned} \frac{d}{dt} q_1 &= G M^{-1} G^T \, p_1 \\ \frac{d}{dt} p_1 &= -\nabla_{q_1} U - K q_1 - \nabla_{q_1} \frac{p_1^T G M^{-1} G^T p_1 + p_2^T B M^{-1} B^T p_2}{2} \end{aligned} \tag{13}$$

and

$$\begin{aligned} \frac{d}{dt} q_2 &= B M^{-1} B^T \, p_2 \\ \frac{d}{dt} p_2 &= -\nabla_{q_2} U - \nabla_{q_2} \frac{p_1^T G M^{-1} G^T p_1 + p_2^T B M^{-1} B^T p_2}{2} \end{aligned} \tag{14}$$

where, for notational convenience, we suppressed the arguments in the mappings $U(q_1, q_2)$, $G(q_1, q_2)$, and $B(q_1, q_2)$.

We are interested in the free energy $\mathcal{H}(q_2, p_2)$ of the system (14). This requires taking the ensemble average in (14) over the variable (q_1, p_1) using the equations of motion (13).[16] In the following section, we show that the potential $U(q_1, q_2)$ does not depend on q_1 in first order approximation. This will greatly simplify the computation of the free energy $\mathcal{H}(q_2, p_2)$ which will be carried out in Section 5.

4. WHY SOFT CONSTRAINTS?

As already mentioned in Section 2, the definition of the soft constraint function g is equivalent to

$$g(q) := K^{-1}\nabla_{\tilde{q}_1}V(q) \tag{15}$$

with $\tilde{q}_1 := \tilde{g}(q)$ and $\tilde{g}$ the hard constraint function. Now, as defined in Section 2, we rewrite the potential energy V as

$$V = U + \frac{g^T K g}{2} \tag{16}$$

and take the gradient w.r.t. $\tilde{q}_1$; i.e.

$$\nabla_{\tilde{q}_1}V = \nabla_{\tilde{q}_1}U + (\frac{\partial q_1}{\partial \tilde{q}_1})^T K g \tag{17}$$

where $q_1 = g(q)$ as before. Now

$$\frac{\partial q_1}{\partial \tilde{q}_1} = I + O(K^{-1}) \tag{18}$$

and, therefore,

$$\nabla_{\tilde{q}_1}V = \nabla_{\tilde{q}_1}U + Kg + O(q_1)\,. \tag{19}$$

This and the definition of g imply that, up to terms of order $O(K^{-1})$, $\nabla_{q_1}U = O(q_1)$. Thus, expanding $U(q_1, q_2)$ in q_1; i.e.

$$U(q_1, q_2) = U(0, q_2) + A(q_2)q_1 + \frac{q_1^T B(q_2) q_1}{2} + \dots\,, \tag{20}$$

we obtain $A(q_2) = O(K^{-1})$ and $B(q_2) = O(1)$. Finally, upon assuming that $K + B(q_2) \approx K$, the potential energy function V can approximately be written as

$$V(q_1, q_2) \approx U(0, q_2) + \frac{q_1^T K q_1}{2} \tag{21}$$

which will be used in the following section to compute the free energy of the torsion dynamics of (1).

Note that, in terms of the hard constraint function $\tilde{q}_1 = \tilde{g}(q)$, the corresponding expansion of the potential energy function $\tilde{U}$ would only yield $\tilde{A}(q_2) = O(1)$ and the approximation (21) could not be applied.

Remark. The modified constraint condition $g(q) = 0$ was also applied by Duan et al.[17] in their modified SHAKE method for constrained energy minimization. The above discussions makes clear why their modification yields improved results over the standard constrained minimization using hard constraints.

5. FREE ENERGY OF TORSION DYNAMICS

In this Section we want to derive an approximation to the free energy $\mathcal{H}(q_2, p_2)$. In a first step we take the ensemble average in (14) over q_1. Neglecting momentum dependent terms and applying the result of Section 4, the Boltzmann distribution function $\rho(q_1)$ is given by

$$\rho(q_1) \approx \frac{1}{C}\exp(-\beta(\frac{q_1^T K q_1}{2}) \;\approx\; \delta(q_1^T q_1) \tag{22}$$

where $\delta(x)$ is Dirac's delta function. Thus averaging over q_1 becomes trivial.

Taking the ensemble average over the momentum p_1 in the expression

$$\nabla_{q_2}\frac{p_1^T G M^{-1} G^T p_1}{2}\,. \tag{23}$$

is a bit more tricky. Using equipartitioning of energy[18] in the kinetic energy, one can show[10] that averaging over p_1 leads to

$$\langle\nabla_{q_2}\frac{p_1^T G(q) M^{-1} G(q)^T p_1}{2}\rangle_{\mathrm{ens}}^{(1)} = \frac{k_B T}{2}\nabla_{q_2}\ln\,\det\,[G(q)M^{-1}G(q)^T] \tag{24}$$

and subsequent averaging over q_1 yields the potential

$$U_F(q_2) = \frac{k_B T}{2}\ln\,\det\,[G(q_2)M^{-1}G(q_2)^T]\,. \tag{25}$$

The potential (25) has been introduced before by Fixman[8] in the context of statistical mechanics. He showed that (25) has to be included to make sure that, in the limit $||K^{-1}|| \to$

0, the unconstraint system (8) and the corresponding constraint system have the same reduced canonical density function in the variable (q_2, p_2). Similar results can be found elsewhere.[7,9,19]

The free energy $\mathcal{H}(q_2, p_2)$ is thus (approximately) given by

$$\mathcal{H}(q_2, p_2) = \frac{p_2^T B(q_2) M^{-1} B(q_2)^T p_2}{2} + U(q_2) + U_F(q_2) \tag{26}$$

or, in terms of the cartesian coordinates $(q, p) \in R^{2n}$, by the Hamiltonian

$$\mathcal{H}(q, p) = \frac{p^T M^{-1} p}{2} + U(q) + U_F(q) + g(q)^T \lambda \tag{27}$$

together with the constraint

$$g(q) = 0\,. \tag{28}$$

The corresponding equations of motion are

$$\begin{aligned} \frac{d}{dt} q &= M^{-1} p \\ \frac{d}{dt} p &= -\nabla U(q) - \nabla U_F(q) - G(q)^T \lambda \\ 0 &= g(q) \end{aligned} \tag{29}$$

Note that ∇U in (29) can be replaced by ∇V and the explicit knowledge of the potential U is therefore not needed.

Finally a few remarks on the practical computation of the soft constraint function g. In many cases it will be possible to split the gradient $\nabla \tilde{U}$ into a strong and a weak contribution. Let us denote the corresponding entries in the potential $\tilde{U}$ by $\tilde{U}_{\text{hard}}$, $\tilde{U}_{\text{soft}}$ respectively. Then g can be simplified to

$$g(q) := \tilde{g}(q) + K^{-1} [\tilde{G}(q) M^{-1} \tilde{G}(q)^T]^{-1} \tilde{G}(q) M^{-1} \nabla \tilde{U}_{\text{hard}}(q)\,. \tag{30}$$

With N the number of particles, the costs for evaluating $\nabla \tilde{U}_{\text{hard}}$ will scale like $O(N)$. This is due to the fact that $\tilde{U}_{\text{hard}}$ will, in general, only include nearest neighborhood interactions; i.e., $\nabla \tilde{U}_{\text{hard}}$ is "banded". In contrast to this the computational costs for $\nabla \tilde{U}_{\text{soft}}$ are of order $O(N^2)$.

Occasionally the term $\nabla_{\tilde{q}_1} \tilde{U}_{\rm hard}$ might become so large that

$$g_i(q)^T K_{ii} g_i(q) \geq k_B T \tag{31}$$

for some i, $1 \leq i \leq m$, where g_i denotes the i-th component of g. This can happen, for example, if close interactions involving Lennard-Jones potentials occur. Since such an strong stretching of the bonds, bending of bond-angles respectively, is non-physical, we suggest to use in practical computations the modified soft constraint function

$$g(q) = \tilde{g}(q) + C \arctan \left(C^{-1} K^{-1} [\tilde{G}(q) M^{-1} \tilde{G}(q)^T]^{-1} \tilde{G}(q) M^{-1} \nabla \tilde{U}_{\rm hard}(q) \right) \tag{32}$$

with

$$C = \frac{1}{\pi} \sqrt{k_B T \, K^{-1}} \,. \tag{33}$$

With this modification the energy in the bond stretching, bond-angle bending respectively, can maximally reach a value of $k_B T/8$.

The matrix $\tilde{G}(q) M^{-1} \tilde{G}(q)^T$ is a banded symmetric matrix with most entries constant. Thus the computation of its inverse requires $O(m)$ operations, m the number of constraints. Often it will be possible to treat the bond stretching by hard constraints and to include only bond-angle bending into the soft constraint function g. This reduces the computational costs by a factor of 2-3.

6. A FORMULATION USING HARD CONSTRAINTS AND A MODIFIED ENERGY FUNCTION

The introduction of the soft constraint function g was very convenient from an analytical point of view. However, the numerical solution of (29) is rather expensive. For example, if one uses a generalization[20] of the well-known SHAKE[11] algorithm to arbitrary constraints, then one obtains

$$
\begin{aligned}
q_{n+1} &= q_n + \Delta t\, M^{-1} p_{n+1/2} \\
p_{n+1/2} &= p_{n-1/2} - \Delta t\, \nabla V(q_n) - \Delta t\, \nabla U_F(q_n) - \Delta t\, G(q_n)^T \lambda_n \\
0 &= g(q_{n+1})
\end{aligned}
\tag{34}
$$

and one has to solve at each integration step a nonlinear system of equations of the form

$$
g(Q + M^{-1} G_n^T \Lambda) = 0\,. \tag{35}
$$

Since evaluation of g requires the computation of $\nabla \tilde{U}_{\rm hard}$, each Newton step would become quite costly. This can be avoided by reformulating (29) as a constraint system on the manifold

$$
\tilde{\mathcal{M}} = \{(q,p) \in \boldsymbol{R}^{2n} : \tilde{g}(q) = 0,\ \tilde{G}(q) M^{-1} p = 0\,\} \tag{36}
$$

which is obtained by setting bond-lengths and bond-angles to their equilibrium values; i.e., $\tilde{g}$ is the hard constraint function. This requires a transformation $w : \tilde{\mathcal{M}} \to \mathcal{M}$ which leaves $q_2 = b(q)$ invariant. In good approximation this transformation is given by

$$
\begin{aligned}
Q &= q + M^{-1} \tilde{G}(q)^T \mu \\
0 &= g(Q)
\end{aligned}
\tag{37}
$$

and $w(q) := Q$. Thus the free energy on $\tilde{\mathcal{M}}$ is defined by

$$
\mathcal{H}(q,p) = \frac{p^T M^{-1} p}{2} + V(w(q)) + \tilde{U}_F(q) + \tilde{g}(q)\lambda \tag{38}
$$

together with the (hard) constraint

$$
\tilde{g}(q) = 0\,. \tag{39}
$$

Here $\tilde{U}_F$ is now given by

$$
\tilde{U}_F(q) = \frac{k_B T}{2} \ln\, \det\, [\tilde{G}(q) M^{-1} \tilde{G}(q)^T]\,. \tag{40}
$$

The corresponding equations of motion are

$$\begin{aligned}
\frac{d}{dt}q &= M^{-1}p \\
\frac{d}{dt}p &= -W(q)^T\nabla V(w(q)) - \nabla\tilde{U}_F(q) - \tilde{G}(q)^T\lambda \\
0 &= \tilde{g}(q)
\end{aligned} \tag{41}$$

with $W(q) = w_q(q)$. These equations can be discretized by the standard SHAKE method; i.e.

$$\begin{aligned}
q_{n+1} &= q_n + \Delta t\, M^{-1}p_{n+1/2} \\
p_{n+1/2} &= p_{n-1/2} - \Delta t\, W(q_n)^T\nabla V(w(q_n)) - \Delta t\,\nabla\tilde{U}_F(q_n) - \Delta t\,\tilde{G}(q_n)^T\lambda_n \\
0 &= \tilde{g}(q_{n+1})
\end{aligned} \tag{42}$$

In principle we have not gained much yet. As defined above, the computation of w and its derivative W is still expensive. However, since $w(q) - q = O(K^{-1})$, the following simplification seems justified: Let us apply one Newton iteration to (37). Upon neglecting terms of order $O(K^{-2})$, the variable Q is now given by

$$Q := q - A(q)\nabla\tilde{U}_{\mathrm{hard}}(q) \tag{43}$$

with the symmetric matrix $A(q)$ defined by

$$A(q) = M^{-1}\tilde{G}(q)^T[\tilde{G}(q)M^{-1}\tilde{G}(q)^T]^{-1}K^{-1}[\tilde{G}(q)M^{-1}\tilde{G}(q)^T]^{-1}\tilde{G}(q)M^{-1}\,.$$

The corresponding mapping $\tilde{w}$ with $\tilde{w}(q) := Q$, $q \in \tilde{\mathcal{M}}$, is $O(K^{-2})$ close to w. Upon using these approximations in (42), the method requires now one additional evaluation of $\nabla\tilde{U}_{\mathrm{hard}}(q)$ per integration step and computation of the derivative of $A(q)\nabla\tilde{U}_{\mathrm{hard}}(q)$ with respect to q. The main computational cost for this are caused by the Hessian of $\tilde{U}(q)$.

The formulation (41) has another advantage. It can easily be implemented into existing methods that use internal coordinates instead of cartesian coordinates (see, for example, Refs. 21 and 22). Since most of these methods use the bond-lengths, bond-angles, and/or torsion-angles as internal degrees of freedom, the constraint $\tilde{g}(q) = 0$ is easily implemented by freezing the bond-lengths and bond-angles to their equilibrium value. The only modification concerns the force field which has now to include the transformation w, $\tilde{w}$ respectively. Of course, one could also write the equations of motion directly in terms of the torsion angles.

7. EXAMPLE

As a numerical example we consider motion of a single butane molecule coupled to a heat bath at $T = 300\,\mathrm{K}$. A united-atom representation of butane is used ($CH_3{-}CH_2{-}CH_2{-}CH_3$). A harmonic potential is used to describe the bond-length fluctuations: $K_r(r-r_0)^2/2$, where r is the actual length of the bond, with $r_0 = 1.53\,\mathring{A}$, and $K_r = 83.7\,\mathrm{kcal/(mol\,\mathring{A}^2)}$.[23] Similarly, a harmonic potential is used for the bond-angle vibrations: $K_\phi(\cos(\phi) - \cos(\phi_0))^2/2$, where ϕ is the actual angle, ϕ_0 is the tetrahedral value of 109.5^o, and the force constant $K_\phi = 43.1\,\mathrm{kcal/mol}$.[23] We do not use Ryckaert-Bellemans potential[24] for the torsion angle. Instead the dihedral interaction is modeled by

$$V_{\mathrm{tors}}(\psi) = \frac{1.5}{2}(1 - \cos(3\psi))\,, \tag{44}$$

where ψ is the torsion angle. A Lennard-Jones potential describes the interaction between the two CH_3 groups as function

$$V_{\mathrm{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{45}$$

with $\epsilon/k_B T = 120\,\mathrm{K}$, $\sigma = 3.2$, and r the distance between the two groups. Of course, we do not claim that these potentials correctly model the dihedral interaction in butane. Here we are only interested in demonstrating the differences between free dynamics and different constraint formulations.

We first determine the effective torsion potential for our model using hard constraints, soft constraints respectively. In other words, we compute

$$V_{\mathrm{eff}}(\psi) = V_{\mathrm{tors}}(\psi) + V_{\mathrm{LJ}}(r(\psi)) \tag{46}$$

where the distance $r(\psi)$ between the two CH_3 groups is either computed by freezing the bond-lengths and bond-angles to their equilibrium value (hard constraints) or by setting the bond-lengths and bond-angles to their values given by $Q = w(q)$ (soft constraints) with w defined by (37). Here we use in (37) the modified soft constraint function (32). The results can be found in Fig. 1. Note the significant difference in the potentials as ψ approaches $\pm\pi$ where the two CH_3 groups get closest. We also compute the effective potential using the simplified $\tilde{w}$. The result compared to the potential obtained for the "exact" w is given in Fig. 2. We also evaluated the Fixman potential (25). It amounts to a few tenths of $k_B T \approx 0.6\,\mathrm{kcal/mol}$ at $T = 300\,\mathrm{K}$ and is therefore small compared to the potential V_{eff}.

To see whether the computed effective potentials reflect the true free energy of the reduced system, we simulate the unconstraint formulation using Langevin dynamics. We compute the distribution function for the torsion angle ψ and compare the result to the distribution functions corresponding to the effective torsion potentials V_{eff} for hard and soft constraints. The results can be found in Fig. 3. Note the excellent agreement between the Boltzmann distribution for the free dynamics and the corresponding distribution function for the constraint dynamics with soft constraints. The distribution function for the constraint formulation using hard constraints exhibits a too low transition-rate from the *trans* to the *cis* conformation; i.e., the constraint system becomes too rigid. This agrees with general observations concerning the application of constraint dynamics (hard constraints) to bond-angle bending.[12]

8. SUMMARY

In this paper, we have derived a new formulation for the torsion dynamics of molecular systems. The resulting equations of motion can either be discretized by the well-known SHAKE method or by methods that use internal coordinates. We have demonstrated for a single butane molecule that our formulation yields a qualitative improvement over standard constraint methods that freeze bond-lengths and bond-angles to their equilibrium value.

9. ACKNOWLEDGMENTS

This work was initiated while the author was visiting Klaus Schulten at the Beckmann Institute in Urbana. The work benefited from discussions with Bernie Brooks (NIH) and Ernst-Walter Knapp (FU Berlin).

REFERENCES

1. Allen, M.P. and Tildesley, D.J., *Computer simulations of liquids*, Oxford University Press, New York, (1987).

2. McCammon, J.A. and Harvey, S.C., *Dynamics of proteins and nucleic acids*, Cambridge University Press, (1987).

3. Brooks III, C.L., Karplus, M., and Pettitt, B.M., *Proteins: A theoretical perspective of dynamics, structure, and thermodynamics*, John Wiley & Sons, New York, (1988).

4. Leimkuhler, B., Reich, S., and Skeel, R.D., Integration methods for molecular dynamics, to appear in *IMA Volumes in Mathematics and its Applications*, Mesirov, J. and Schulten, K., editors, Springer Verlag, (1995).

5. Verlet, L., Computer experiments on classical fluids, Part I, *Phys. Rev.*, **159**, 98–103, (1967).

6. Rubin, H. and Ungar, P., Motion under a strong constraining force, *Comm. Pure Appl. Math.*, **10**, 65–87, (1957).

7. Pear, M.R. and Weiner, J.H., Brownian dynamics study of a polymer chain of linked rigid bodies, *J. Chem. Phys.*, **71**, 212–224, (1979).

8. Fixman, M., Classical statistical mechanics of constraints: A theorem and applications to polymers, *Proc. Nat. Acad. Sci.*, **71**, 3050–53, (1974).

9. van Kampen, N.G. and Lodder, J.J., Constraints, *Am. J. Phys.*, **52**, 419–424, (1984).

10. Reich, S., Smoothed Dynamics of Highly Oscillatory Hamiltonian Systems, *Physica D*, to appear, (1995).

11. Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C., Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *J. Comput. Phys.*, **23**, 327–342, (1977).

12. van Gunsteren, W.F. and Karplus, M., Effects of constraints on the dynamics of macromolecules, *Macromolecules*, **15**, 1528–44, (1982).

13. Brooks, B.R., Zhou, J., and Reich, S., Elastic molecular dynamics with flexible constraints, submitted.

14. Peskin, C.S. and Schlick, T., Molecular dynamics by backward Euler method, *Comm. Pure and Appl. Math.*, **42**, 1001-1031, 1989.

15. Boothby, W.M., *An introduction to differentiable manifolds and Riemannian geometry*, 2nd edition, Academic Press, (1986).

16. Kuczera, K., Conformational free energy from thermodynamic integration simulations, Techn. Report, University of Kansas, Lawrence, (1995).

17. Duan, Y., Kumar, S., Rosenberg, J.M., and Kollman, P.A., Gradient SHAKE: An improved method for constrained energy minimization in macromolecular simulations, *J. Comp. Chem.* **16**, 1351–1356, 1995.

18. Pathria, P.K., *Statistical mechanics*, Pergamon Press, Oxford, (1972).

19. Helfand, E., Flexible vs. rigid constraints in statistical mechanics, *J. Chem. Phys.*, **71**, 5000–07, (1979).

20. Reich, S., Symplectic Integration of Constrained Hamiltonian Systems by Composition Methods, *SIAM J. Numer. Anal.*, to appear, (1996).

21. Jain, A., Unified formulation of dynamics for serial rigid multibody systems, *Journal of Guidance, Control, and Dynamics*, **14**, 531–542, (1991).

22. Mazur, A.K., Dorofeyev, V.E., and Abagyan, R.A., Derivation and Testing of Explicit Equations of Motion for Polymers Described by Internal Coordinates, *J. Comp. Phys.*, **92**, 261–272, (1991).

23. Rosenberg, R.O., Berne, B.J., and Chandler, D., Isomerization Dynamics in Liquids by

Molecular Dynamics, *Chem. Phys. Lett.*, **75**, 162, (1980).

24. Ryckaert, J.P. and Belemans, A., Molecular dynamics of liquid N-butane near its boiling point, *Chem. Phys. Lett.*, **30**, 123, (1975).

FIGURES

Fig. 1. Effective torsion potential in kcal/(mol rad) for butane using hard constraints (dotted line) versus the potential obtained with soft constraints (solid line).

Fig. 2. Effective torsion potential in kcal/(mol rad) using soft constraints given by the "exact" transformation w (solid line) and by the simplified $\tilde{w}$ (dotted line).

Fig. 3. Boltzmann distribution function for torsion angle using hard constraints (dash-dotted line), soft constraints (solid line), and unconstrained formulation (dotted line).