

## ARTICLES

## Constant Temperature Constrained Molecular Dynamics: The Newton–Euler Inverse Mass Operator Method

Nagarajan Vaidehi,<sup>†</sup> Abhinandan Jain,<sup>‡</sup> and William A. Goddard, III<sup>\*,†</sup>*Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering (CN 9114), California Institute of Technology, Pasadena, California 91125, and Jet Propulsion Laboratory/California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109**Received: October 13, 1995; In Final Form: April 13, 1996*<sup>⊗</sup>

The Newton–Euler inverse mass operator (NEIMO) method for internal coordinate molecular dynamics (MD) of macromolecules (proteins and polymers) leads to stable dynamics for time steps about 10 times larger than conventional dynamics (e.g., 20 or 30 fs rather than 1 or 2 fs for systems containing hydrogens). NEIMO is practical for large systems since the computation time scales linearly with the number of degrees of freedom  $\mathcal{N}$  (instead of the  $\mathcal{N}^3$  scaling for conventional constrained MD methods). In this paper we generalize the NEIMO formalism to the Nosé (and Hoover) thermostat to derive the Nosé and Hoover equations of motion for constrained canonical ensemble molecular dynamics. We also examined the optimum mass,  $Q$ , determining the time scale ( $\tau_s$ ) for exchange of energy with the heat bath for NEIMO–Hoover dynamics of polymers. We carried out NEIMO–Hoover simulations on the amorphous polymers poly(vinyl chloride) and poly(vinylidene fluoride), where we find that time steps of 20–30 fs lead to stable dynamics (10 times larger than for Cartesian dynamics). The computational efficiency of the NEIMO canonical MD method should make it a powerful tool for MD simulations of macromolecular materials.

## 1.0. Introduction

For studies of the conformations and dynamics of polymers and proteins, it is often useful to simplify the description of the systems by constraining such structural properties as bond lengths and bond angles so that the focus can be on the dihedral angles distinguishing the conformations.<sup>1–4</sup> With such constraints the number of degrees of freedom (dof) drops from  $3N$  (where  $N$  is the number of atoms) to  $\mathcal{N}$ . Thus, for a polyethylene polymer,  $C_pH_{2p+2}$ ,  $3N = 9p + 6$ , while the number of torsional dof is  $\mathcal{N} = p - 1$ . In addition to simplifying the analysis, such constrained molecular dynamics (MD) can allow significantly increased time steps.

With constrained MD, Newton's equation of motion becomes

$$\mathcal{M}(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + \mathbf{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) = \mathcal{T}(\boldsymbol{\theta}) \quad (1)$$

where  $\boldsymbol{\theta}$  denotes the vector of the generalized coordinates (e.g. torsional angles),  $\mathcal{T}$  denotes the vector of generalized forces (e.g., torques),  $\mathcal{M}$  denotes the mass matrix (moment of inertia tensor), and  $\mathbf{C}$  includes the Coriolis forces. The dynamics of motion is obtained by solving (1) for the acceleration

$$\ddot{\boldsymbol{\theta}} = \mathcal{M}^{-1}[\mathcal{T}(\boldsymbol{\theta}) - \mathbf{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})] \quad (2)$$

and integrating to obtain new velocities and coordinates. The problem here is that  $\mathcal{M}$  is a  $\mathcal{N} \times \mathcal{N}$  matrix, and hence the calculation of  $\mathcal{M}^{-1}$  in (2) involves a computational cost scaling as  $\mathcal{N}^3$ . Since the other parts of the calculation generally scale linearly in  $\mathcal{N}$ , calculation of  $\mathcal{M}^{-1}$  can become the dominant

cost in the dynamics of large systems. Thus for 1001 atoms ( $p = 333$ ), it is necessary to invert a  $332 \times 332$  matrix every step of the dynamics, costing more than the rest of the dynamics. Using the cell multipole method,<sup>5</sup> MD has been demonstrated<sup>4</sup> to be practical for  $10^6$  atoms ( $p = 333\,333$ ). However, inverting the  $333\,332 \times 333\,332$  matrix would be clearly impractical.

To solve this problem we use the Newton–Euler inverse mass operator (NEIMO) approach to calculate the  $\ddot{\boldsymbol{\theta}}$  of (2) directly *without* going through the step of explicitly calculating  $\mathcal{M}$  and then inverting. NEIMO is based on a spatial operator algebra formalism and *scales linearly* with  $\mathcal{N}$ . The application of NEIMO to (standard microcanonical) dynamics of proteins, including the Tomato Bushy Stunt virus, has been reported elsewhere.<sup>4</sup>

Standard Newtonian dynamics leads to conservation of the total energy along the trajectory. Thus the collection of points from the MD describes a *microcanonical* ensemble (in which the energy  $E$ , volume  $V$ , and number of particles  $N$  are conserved). However, normal experimental conditions have the system in contact with a heat bath with constant temperature. Thus we want the MD simulation to simulate a *canonical ensemble* appropriate for describing  $(T, V, N)$  and  $(T, P, N)$  canonical ensembles. This extension was carried out by Nosé<sup>6</sup> and Hoover<sup>7,8</sup> for Cartesian dynamics.

In this paper we derive the theory of constrained dynamics in the  $(T, V, N)$  canonical ensemble and present a NEIMO algorithm (based on spatial operator algebra and the corresponding recursive equations) to solve the NEIMO–Nosé and NEIMO–Hoover equations.

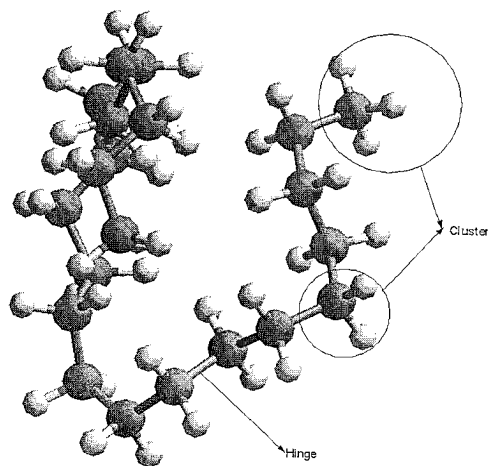
Section 2 contains a brief overview of the microcanonical NEIMO method. Section 3 contains the derivation of Nosé and Hoover equations of motion in internal coordinates. Section 4

\* To whom correspondence should be addressed.

<sup>†</sup> Beckman Institute (wag@wag.caltech.edu).

<sup>‡</sup> Jet Propulsion Laboratory.

<sup>⊗</sup> Abstract published in *Advance ACS Abstracts*, June 1, 1996.



**Figure 1.** Clusters and hinges for the NEIMO–Hoover simulations of polyethylene oligomers. The large atoms are carbon, and the small ones are hydrogen.

contains the recursive algorithm for solving the NEIMO–Hoover equations of motion. In section 5 we discuss the optimization of the Nosé mass parameter for NEIMO–Hoover simulations (with bond lengths and bond angles constrained). Illustrative applications of the NEIMO–Hoover method for amorphous polymers and other polymers are shown in section 6.

## 2.0. NEIMO Method in the Microcanonical Ensemble

The NEIMO method offers a  $\mathcal{AM}$  algorithm for solving the equations of motion, (1), for constant energy dynamics. The system is partitioned in terms of rigid bodies termed *clusters* which are connected by *hinges* each of which can have one to six dof. The cluster is a group of atoms moving as a rigid unit in a molecule. It can consist of a single atom, a methylene group, a phenyl ring, an  $\alpha$  helix, or even an entire domain of a protein. The hinge describes the relative motion between two adjacent clusters. A typical application to polymers, proteins, or DNA is to freeze all bonds and angles. In this case the hinge becomes one-dimensional, describing a torsion. If only the distance between two clusters is fixed, the hinge is five-dimensional. Thus a molecule consists of one or more clusters with hinges connecting adjacent clusters. The union of clusters connected to each other by hinges is referred to as a *chain*. A chain can be cross-linked (multiply connected) or open. A system may consist of single or multiple chains. For each chain there is a designated *base cluster* which is connected to the reference coordinate system by a hinge having the full six dof. This enables each chain to have full freedom to translate and orient in space.

Within a chain the relationship between adjacent clusters is described in terms of *parent* and *child* clusters. The base cluster has no parents but can have children clusters; it is the parent cluster of each of these children. Each of these children clusters may in turn have zero, one, or more children, branching outward from the base. Outward branching can continue in a topological tree with each cluster having zero, one, or more children, but each child having only one parent cluster. Clusters with no children are called *tips*. A chain can be infinite, but this modifies the algorithm, which will be discussed elsewhere.

In a *serial chain* (such as a linear polymer), there is a single tip cluster and each cluster between the base and the tip has a unique parent and unique child. These concepts are illustrated in Figure 1.

The NEIMO methodology has been developed for solving the dynamical equations of motion for serial chains, for tree

topologies, and for closed loop systems.<sup>2,9,10</sup> For each cluster, the spatial quantities (position, velocity, acceleration, and momentum) are calculated recursively using the spatial quantities of the parent cluster. The derivation of the NEIMO algorithm considers the spatial quantities for all the clusters to be stacked in a matrix called the spatial operator. The spatial operator corresponding to the mass matrix is then factorized using innovations operator factorization<sup>9,10</sup> consisting of square and invertible factors. The NEIMO algorithm involves no explicit calculation or inversion of the mass matrix. There are three steps to solving the equations of motion:

1. *Velocities.* A base to tips recursion, during which the spatial velocity ( $\dot{\theta}$ ) of each cluster is calculated from the motion of its associated hinge as well as the motion of its parent cluster.

2. *Forces.* A tips to base recursion, during which the effective force ( $\mathcal{T} - \mathbf{C}$ ) acting on each cluster is derived from the direct force terms acting on the cluster (Cartesian forces, hinge torques, Coriolis forces, and other velocity dependent forces) plus the indirect forces due to its children. This step requires a knowledge of the velocities in order to calculate the Coriolis forces. Other quantities<sup>2</sup> related to the inverse of the mass matrix are also calculated during this recursion.

3. *Accelerations.* A final base to tips recursion to compute the accelerations ( $\ddot{\theta}$ ) for all clusters. The details of this algorithm along with the definition of the spatial operators is described in ref 2.

The NEIMO algorithm in the microcanonical ensemble was combined<sup>4</sup> with the POLYGRAF MD package<sup>11</sup> and successfully applied for various polypeptides and proteins.<sup>4</sup> NEIMO simulations on polypeptides show that time steps as large as 20–30 fs can be used, even for systems containing explicit hydrogens. The NEIMO method in the microcanonical ensemble has also been applied to the MD simulations of amorphous polymers (e.g., poly(vinyl chloride)) described using periodic boundary conditions where we find that time steps of 20 fs generally lead to stable dynamics.

## 3.0. Canonical Ensemble NEIMO Method

Several methods have been developed<sup>12–14</sup> to simulate constant temperature systems. Constant kinetic energy methods scale the velocities<sup>13,15</sup> or momenta<sup>16</sup> so that the total kinetic energy remains constant. However, this leads to incorrect fluctuations in kinetic energy and hence does not produce configurations corresponding to a canonical distribution in momentum space. Andersen<sup>17</sup> used discontinuous stochastic collisions to induce the canonical distribution (essentially MD with the velocities of the particles governed stochastically).

Nosé<sup>6</sup> made a major advance by

- i. extending the system to contain an additional degree of freedom representing the interaction of the system with a heat bath and

- ii. choosing the form of the Hamiltonian so that the (equilibrium) dynamics generates the proper canonical distribution in both momentum and configuration space. In the Nosé *extended* system the physical system is placed in an external bath with one additional coordinate variable  $s$  (corresponding to a time-scaling variable) and its conjugate momentum  $p_s$ . The equations of motion for  $s$  drive the changes of velocities with time which in turn drive the changes in the thermal energies with time. This leads to the proper canonical description of the thermal energy. However, the real time interval corresponding to each time step is unequal, which complicates calculation of dynamical quantities involving Fourier transforms. Hoover<sup>7,8</sup> proposed an alternative formulation in which the real time is explicit. Since the Hoover formulation involves real

velocities and real time steps, it allows a simpler calculation of dynamical properties (fast Fourier transforms).

We now derive the Nosé and Hoover equations of motion in the internal coordinate dynamics framework. We also derive a solution to these equations of motion on the basis of the recursive algorithm previously derived for the NEIMO micro-canonical ensemble.<sup>2</sup> The computational time for this recursive algorithm also scales linearly with  $\mathcal{N}$ .

### 3.1. Derivation of NEIMO–Nosé Equations of Motion.

The equations of motion for the Nosé extended system in the constrained internal coordinate framework are derived using the Lagrangian formalism. Equations of motion are derived for a system with clusters connected by hinges, wherein each hinge can have one to six dof. These equations of motion are general and can be used with any tree topology or any number and size of clusters and hinges. We first derive the Nosé type of equations of motion and then extend it to the Hoover formalism.

Consider a system with generalized coordinates ( $\theta$ ) and  $\mathcal{N}$  generalized velocities ( $\dot{\theta}$ ). We represent the virtual Nosé variables with a prime and the real variables without primes (this is opposite the convention of ref 6). The system is placed in a bath of temperature  $T_B$ , with the bath variables being  $s$  and  $p_s$ . In the Nosé formulation the real time  $t$  is related to the virtual time  $t_s$  by

$$\frac{d}{dt} = s \frac{d}{dt_s} \quad (3)$$

Thus the virtual Nosé generalized variables are defined by

$$\theta_i = \theta'_i \quad (4a)$$

and

$$\dot{\theta}_i = s \dot{\theta}'_i \quad (4b)$$

The Lagrangian for the extended system in Nosé variables (i.e., virtual variables) is given by

$$\mathcal{L} = \text{KE} - \text{PE} = \frac{1}{2} s^2 \sum_{i,j=1}^{\mathcal{N}} \dot{\theta}'_i M_{ij}(\theta') \dot{\theta}'_j - \Phi(\theta') + \frac{1}{2} Q (s')^2 - gkT_B \ln s \quad (5a)$$

where it is understood that  $s$  is in the extended system, KE is the kinetic energy, and PE is the potential energy of the extended system. The  $M_{ij}$  in (5a) are the elements of the mass matrix, and  $\Phi$  is the potential energy of the physical system (including the potential energy due to van der Waals and Coulomb forces along with the torsional internal energy). The first and the second terms of (5a) correspond to the kinetic energy and the potential energy of the physical system in internal coordinates. The third and the fourth terms on the right-hand side of (5a) represent the kinetic and the potential energy corresponding to the bath variable  $s$  where  $T_B$  is the temperature of the external bath and  $k$  is the Boltzmann constant. Nosé showed<sup>12</sup> that choosing

$$g = \mathcal{N} + 1 \quad (5b)$$

leads to a partition function of the extended system which when properly integrated with respect to the bath variables leads to the canonical form for the partition function of the physical system as

$$Z_{\mathcal{N}} = \langle e^{-H_0/kT_B} \rangle$$

Thus the trajectory of the extended system leads to a canonical distribution of the coordinates and momenta of the physical system.

From (5a) the conjugate momentum is given by

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}'_k} = s^2 \sum_j M_{kj} \dot{\theta}'_j \quad (6)$$

(where we used  $M_{jk} = M_{kj}$ ), and the equation for the conjugate position is

$$\frac{\partial \mathcal{L}}{\partial \theta'_k} = \frac{1}{2} s^2 \sum_{ij} \dot{\theta}'_i \frac{\partial M_{ij}}{\partial \theta'_k} \dot{\theta}'_j - \frac{\partial \Phi}{\partial \theta'_k} \quad (7)$$

In Nosé variables the Lagrangian equations of motion are given by

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\theta}'_k} \right) - \frac{\partial \mathcal{L}}{\partial \theta'_k} = 0 \quad (8)$$

Substituting (6) and (7) into (8) leads to

$$\frac{d}{dt} (s^2 \sum_j M_{kj} \dot{\theta}'_j) - \frac{1}{2} s^2 \sum_{ij} \dot{\theta}'_i \frac{\partial M_{ij}}{\partial \theta'_k} \dot{\theta}'_j + \frac{\partial \Phi}{\partial \theta'_k} = 0 \quad (9)$$

Although the contribution to  $\partial \Phi / \partial \theta_k$  due to internal coordinates is straightforward, the contributions due to nonbond interactions are tedious to calculate.<sup>1</sup> Consequently, we rewrite the potential energy as

$$\Phi = \Phi_{\text{int}} + \Phi_{\text{NB}} \quad (10)$$

where  $\Phi_{\text{int}}$  is the contribution due to internal coordinates (e.g., torsions) and  $\Phi_{\text{NB}}$  contains the Coulomb, van der Waals, and external forces. The gradient is given by

$$\frac{\partial \Phi}{\partial \theta_k} = \frac{\partial \Phi_{\text{int}}}{\partial \theta_k} + \sum_{\alpha} \frac{\partial \Phi_{\text{NB}}}{\partial X_{\alpha}} \frac{\partial X_{\alpha}}{\partial \theta_k} \quad (11)$$

where the summation is over the  $3N$  Cartesian coordinates. The second term which is the gradient of  $\Phi_{\text{NB}}$  with respect to internal coordinates is computationally tedious. Substituting (11) into (9) and rearranging leads to

$$\sum_j M_{kj} \ddot{\theta}'_j + \mathbf{C}_N(\theta, \dot{\theta}') + \mathbf{F}_N(\theta, \dot{\theta}') = \mathcal{F}_N(\theta) \quad (12)$$

where

$$\mathbf{C}_N(\theta, \dot{\theta}') = \sum_j \dot{M}_{kj} \dot{\theta}'_j - \frac{1}{2} \sum_{ij} \dot{\theta}'_i \frac{\partial M_{ij}}{\partial \dot{\theta}'_k} \dot{\theta}'_j + \frac{1}{s^2} \sum_{\alpha} \frac{\partial \Phi_{\text{NB}}}{\partial X_{\alpha}} \frac{\partial X_{\alpha}}{\partial \theta_k} \quad (13)$$

$$\mathbf{F}_N(\theta, \dot{\theta}') = \frac{2}{s} \frac{ds}{dt_s} \sum_j M_{kj} \dot{\theta}'_j \quad (14)$$

and

$$\mathcal{F}_N(\theta) = -\frac{1}{s^2} \frac{\partial \Phi_{\text{int}}}{\partial \theta'_k} = -\frac{1}{s^2} \frac{\partial \Phi_{\text{int}}}{\partial \theta_k} \quad (15)$$

$\mathbf{F}_N(\theta, \dot{\theta}')$  corresponds to a velocity-dependent friction-like force and could be included in the Coriolis term  $\mathbf{C}_N(\theta, \dot{\theta}')$ . The Coriolis term also contains the nonbond and the external forces. The NEIMO algorithm has the distinct advantage of not requiring the gradient of the nonbond and the external potential

with respect to internal coordinates. Instead it can handle the gradient of the nonbond and external potentials with respect to the Cartesian coordinates which is computationally advantageous. In the NEIMO algorithm the forces can be written as

$$\frac{\partial \Phi_{\text{int}}}{\partial \theta_k} + \sum_{\alpha} \frac{\partial \Phi_{\text{NB}}}{\partial X_{\alpha}}$$

(12) is similar to the microcanonical equations given by (1). The major differences are the scaling of the generalized forces by the variable  $s$  and the additional friction-like term. The Lagrangian equation of motion for  $s$  is

$$\frac{d}{dt}(Q\dot{s}) - s \sum_{ij} \dot{\theta}'_i M_{ij} \dot{\theta}'_j + \frac{gkT_B}{s} = 0$$

or

$$\ddot{s} = \frac{1}{Qs} [s^2 \sum_{ij} \dot{\theta}'_i M_{ij} \dot{\theta}'_j - gkT_B] \quad (16)$$

The kinetic energy of the physical system is

$$\text{KE} = \frac{1}{2} s^2 \sum_{ij} \dot{\theta}'_i M_{ij} (\dot{\theta}')_j = \frac{1}{2} \mathcal{N} kT \quad (17)$$

where  $T$  defines an instantaneous temperature. Thus, (16) becomes

$$\ddot{s} = \frac{1}{Qs} [\mathcal{N} kT - gkT_B] \quad (18)$$

We prefer to write  $Q$  in terms of the time scale of relaxation of the bath variables as<sup>18,19</sup>

$$\tau_s^2 = \frac{Q}{\mathcal{N} kT_B} \quad (19)$$

Substituting for  $Q$  from (19) into (18) leads to

$$\ddot{s} = \frac{1}{s\tau_s^2} \left[ \frac{T}{T_B} - \frac{g}{\mathcal{N}} \right] \quad (20)$$

Substituting for  $g$  from (5b) leads to (21)

$$\ddot{s} = \frac{1}{s\tau_s^2} \left[ \left( \frac{T}{T_B} - 1 \right) - \frac{1}{\mathcal{N}} \right] \quad (21)$$

as the equation of motion for  $s$ . (12) and (21) are the fundamental equations to be solved for the NEIMO–Nosé dynamics.

### 3.2. NEIMO–Hoover Dynamics Equations of Motion.

The Nosé equations of motion involve virtual time, which implies unequal time steps in real time. This is inconvenient for the analysis of dynamical properties (e.g., fast Fourier transforms), and Hoover<sup>8</sup> proposed using (3) to transform the Nosé equations of motion into real variables

$$\frac{dA}{dt_s} = \frac{1}{s} \frac{dA}{dt} \quad (22)$$

Substituting for the Nosé variables in (12) leads to

$$\sum_j M_{kj} \ddot{\theta}_j + \mathbf{C}_H(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) + \mathbf{F}_H(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) = \mathcal{T}_H \quad (23)$$

where

$$\mathbf{C}_H(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) = \sum_j \dot{M}_{kj} \dot{\theta}_j - \frac{1}{2} \sum_{ij} \dot{\theta}'_i \frac{\partial M_{ij}}{\partial \theta_k} \dot{\theta}_j + \sum_{\alpha} \frac{\partial \Phi_{\text{NB}}}{\partial X_{\alpha}} \quad (24)$$

$$\mathbf{F}_H(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) = \zeta \sum_j M_{kj} \dot{\theta}_j \quad (25)$$

and

$$\mathcal{T}_H(\boldsymbol{\theta}) = - \frac{\partial \Phi_{\text{int}}}{\partial \theta_k} \quad (26)$$

Thus the NEIMO–Hoover equation, (23), has the same form as (1) and (12). With NEIMO–Hoover the additional frictional term due to the canonical ensemble is  $\zeta \sum_j M_{ij} \dot{\theta}_i$ , which is included in the Coriolis force terms.  $\mathcal{T}_H(\boldsymbol{\theta})$  is the vector of generalized forces.

Transforming (20) to real variables leads to

$$\dot{\zeta} = \frac{1}{\tau_s^2} \left[ \frac{T}{T_B} - \frac{g}{\mathcal{N}} \right] \quad (27)$$

where

$$\zeta = \frac{1}{s} \frac{ds}{dt_s} \quad (28)$$

From (25) we see that  $\zeta$  is a friction coefficient. Hoover showed using the Liouville equation that the probability density in conserved only if

$$g = \mathcal{N}$$

in (27). This leads to

$$\dot{\zeta} = \frac{1}{\tau_s^2} \left[ \frac{T}{T_B} - 1 \right] \quad (29)$$

(23) and (29) are the fundamental equations for NEIMO–Hoover dynamics. These equations are solved for each cluster in the tree topology using the recursive algorithm described in the next section.

### 4.0. The NEIMO–Hoover Equations of Motion: Recursive Solution

Since the NEIMO–Hoover equations of motion (23) have the same form as (1), all the spatial operator equations and factorization discussed in ref 2 hold for the solution of these equations. The three steps described in section 2 for the recursive NEIMO algorithm are also the major steps used in the solution of the NEIMO–Hoover equations of motion.

The *innovations operator factorization*<sup>9,10</sup> provides a closed form expression for the block  $\mathbf{LDL}^T$  decomposition of the mass matrix  $\mathcal{M}^{-1}$ . The operator expression for the mass matrix inverse in the innovations operator factorization is given by<sup>2</sup>

$$\mathcal{M}^{-1} = [\mathbf{I} - \mathbf{H}\boldsymbol{\psi}\mathbf{K}]^T \mathcal{D}^{-1} [\mathbf{I} - \mathbf{H}\boldsymbol{\psi}\mathbf{K}] \quad (30)$$

The factor  $[\mathbf{I} - \mathbf{H}\boldsymbol{\psi}\mathbf{K}]$  is square, block lower triangular, and nonsingular.  $\mathbf{H}$  is the hinge matrix representing the relative motion characteristics of the  $m$  dof of each hinge ( $m = 1$  for torsions only);  $\boldsymbol{\psi}$  is a lower diagonal spatial transformation matrix.  $\mathbf{K}$  is a spatial operator defined<sup>2</sup> in terms of the body inertia and the hinge characteristics. The matrix  $\mathcal{D}$  is block

diagonal with  $\mathcal{N} \times m \times m$  subblocks ( $1 \times 1$  for torsion only); the transpose in (30) is indicated by a superscript T.

For normal NEIMO dynamics (microcanonical ensemble) the accelerations are obtained by solving the equation

$$\ddot{\theta} = \mathcal{M}^{-1}[\mathcal{T} - \mathbf{C}] \quad (31)$$

From (23) we see that for NEIMO–Hoover canonical dynamics the accelerations become

$$\ddot{\theta} = \mathcal{M}^{-1}[\mathcal{T} - \mathbf{C}_H - \zeta \mathcal{M} \dot{\theta}] \quad (32)$$

or

$$\ddot{\theta} = \mathcal{M}^{-1}[\mathcal{T} - \mathbf{C}_H] - \zeta \dot{\theta} \quad (33)$$

where  $-\zeta \dot{\theta}$  is the friction-like correction term to the (microcanonical) accelerations. The operator expression for  $\ddot{\theta}$  [obtained by substituting the operator expression for  $\mathcal{M}^{-1}$  from (30) into (33)] is given by

$$\ddot{\theta} = [\mathbf{I} - \mathbf{H}\psi\mathbf{K}]^T \mathcal{D}^{-1}[\mathcal{T} - \mathbf{H}\psi(\mathbf{K}\mathcal{T} + \mathbf{P}\mathbf{a} + \mathbf{b} + \hat{\mathbf{f}}_c)] - \mathbf{K}^T \psi^T \mathbf{a} - \zeta \dot{\theta} \quad (34)$$

where  $\mathbf{P}$  is the articulated poor inertia<sup>2</sup>, while  $\mathbf{a}$  and  $\mathbf{b}$  are the Coriolis and gyroscopic forces.

Equation 34 is built up from the following sequence of expressions:

$$\mathbf{z} = \psi[\mathbf{K}\mathcal{T} + \mathbf{P}\mathbf{a} + \mathbf{b} + \hat{\mathbf{f}}_c]$$

$$\boldsymbol{\epsilon} = \mathcal{T} - \mathbf{H}\mathbf{z}$$

$$\boldsymbol{\nu} = \mathcal{D}^{-1}\boldsymbol{\epsilon}$$

$$\boldsymbol{\alpha} = \psi[\mathbf{H}^T \boldsymbol{\nu} + \mathbf{a}]$$

$$\ddot{\theta} = \boldsymbol{\nu} - \mathbf{K}^T \boldsymbol{\alpha} - \zeta \dot{\theta}$$

Each of these above expressions has been shown<sup>9,10</sup> to be of order  $\mathcal{N}$ . From (33) and (34) we see that the only difference in the NEIMO–Hoover equations of motion is the additional computation of the friction-like term,  $\zeta \dot{\theta}$ .

The leapfrog Verlet algorithm is used to obtain velocities and coordinates from the accelerations. Since the NEIMO algorithm requires a knowledge of the velocities at the current step (to compute the Coriolis forces, see (24)) in order to calculate the accelerations, we modified the Verlet algorithm as follows:

- i. We estimate the velocity at the current step  $n$  as

$$\dot{\theta}_n = 1.5\dot{\theta}_{n-(1/2)} - 0.5\dot{\theta}_{n-(3/2)}$$

- ii. We use  $\dot{\theta}_n$  in the recursive solutions to the NEIMO–Hoover equations to obtain the accelerations  $\ddot{\theta}_n$ .

- iii. Using  $\ddot{\theta}_n$  we obtain the velocity  $\dot{\theta}_{n+(1/2)}$ . This is used to re-estimate  $\dot{\theta}_n$

$$\dot{\theta}_n = 0.5\dot{\theta}_{n-(1/2)} + 0.5\dot{\theta}_{n+(1/2)}$$

- iv. Steps ii and iii are repeated until convergence is reached. The criterion used for convergence is that the difference in the velocity is less than 0.001 MD units (length = Å, time = 0.0488 ps). This convergence is generally reached after one or two iterations, adding negligibly to the computational costs. The Hoover dynamical variable  $\zeta$  is also integrated using the Verlet algorithm,

$$\zeta_{n+1} = \zeta_n + \delta D_{n+(1/2)} \quad (35)$$

where  $\zeta_n$  is the value of the friction coefficient at the step  $n$  and  $\delta$  is the time step used.  $D_{n+(1/2)}$  at the half-step is given by

$$D_{n+(1/2)} = \frac{1}{\tau_s^2} \left[ \frac{T_{n+(1/2)}}{T_B} - 1 \right] \quad (36)$$

where  $T_{n+(1/2)}$  is the temperature of the system at the  $n + 1/2$  step. The computer code for the NEIMO–Hoover simulations was interfaced with POLYGRAF<sup>11</sup> for carrying out calculations reported here.

## 5.0. Optimization of the Nosé Mass Parameter

Although the mass parameter  $Q$  does not affect the canonical distribution, it affects the rate at which the system attains equilibrium. Various previous studies for Cartesian molecular dynamics of Lennard-Jones systems (Nosé,<sup>18</sup> DiTolla and Ronchetti,<sup>20</sup> and Cho and Joannopoulos<sup>21</sup>) have shown that a broad range of values for  $Q$  (about 2 orders of magnitude) leads to similar properties, but too small or too large a value of  $Q$  leads to unstable dynamics and different properties.

We examined the appropriate range of values of  $Q$  for NEIMO–Hoover dynamics of polymers (with the bond angles and bond lengths constrained). This is the first test for the best  $Q$  for molecular systems containing internal valence forces. We will express the heat bath mass parameter  $Q$ , in terms of the time constant  $\tau_s$  given by (19), where  $\tau_s$  is the relaxation time of the Nosé variable  $s$ .

The lower limit on the value of  $\tau_s$  is placed by the time step of integration ( $\delta$ ). For stable dynamics one would expect that  $\delta$  is sufficiently small that there be at least 10 steps in the period corresponding to  $\tau_s$ , leading to

$$10\delta \leq 2\pi\tau_s \quad (37)$$

or

$$\tau_s \geq \frac{10}{2\pi}\delta = 1.6\delta \quad (38)$$

There are two time scales<sup>18,20,21</sup> involved in the relaxation toward equilibrium of the extended system.

- i. The relaxation of interatomic forces depends on the fundamental frequencies of the molecular system and the coupling between these modes.

- ii. The relaxation of the heat bath variable ( $s$ ) to achieve thermal equilibrium leads to additional forces depending on  $\tau_s$ .

An estimate of the total simulation time can be obtained by considering the long time fluctuations in the  $\langle s \rangle$ . Applying the equilibrium condition  $\langle T \rangle = T_B$  to (21) and assuming  $\langle s \rangle \approx 1$  leads to

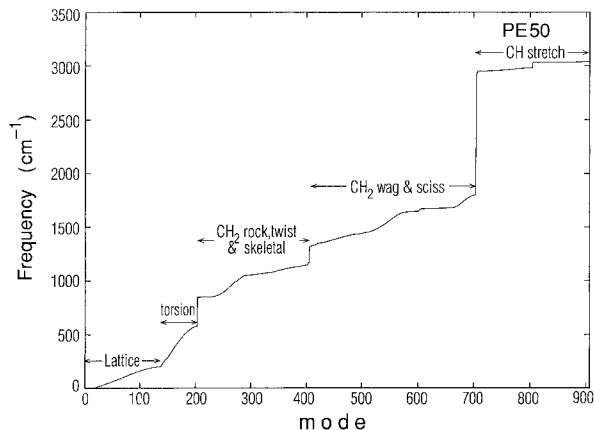
$$\langle \dot{s} \rangle = -\frac{1}{\mathcal{N}\tau_s^2} \quad (39)$$

This suggests a harmonic behavior for the long time fluctuations in  $\langle s \rangle$ , leading to a characteristic relaxation time of

$$\tau_{\text{slong}} = \sqrt{\mathcal{N}}\tau_s \quad (40)$$

the period being  $2\pi$  longer. In order to obtain good averaging, the total time of the simulations should be at least 20 times this period, leading to

$$t_{\text{total}} = 20(2\pi\tau_{\text{slong}}) = 40\pi\sqrt{\mathcal{N}}\tau_s \quad (41)$$



**Figure 2.** Vibrational frequencies of pe50 from the velocity autocorrelation function; 400 ps of Cartesian dynamics ( $\tau_s = .01$  ps). These calculations indicate that the torsion frequencies are in the range of 200–600  $\text{cm}^{-1}$ .

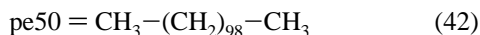
Thus the above analysis suggests that in order to obtain good dynamical behavior

i. From (38)  $\tau_s$  should be larger than the time step for integration which in turn must be small compared to the period of interatomic motions.

ii. From (41) the total simulation time should exceed  $40\pi(\mathcal{N})^{1/2}\tau_s$ .

We now analyze the results of the NEIMO–Hoover simulations for molecular systems using various values of  $\tau_s$  to check the validity of (38) and (41). The choice of  $\tau_s$  is dependent on how fast the system attains equilibrium and the ergodicity of the system. Constrained MD simulations with all bond lengths and bond angles fixed (using only torsional dof) were carried out using the NEIMO–Hoover algorithm described in section 4.

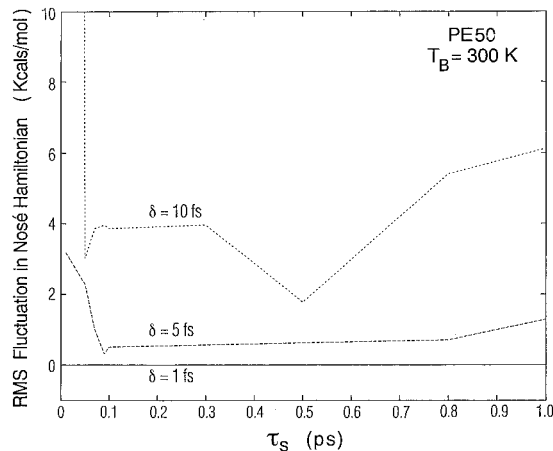
**5.1. Tests on pe50.** We consider the polymer



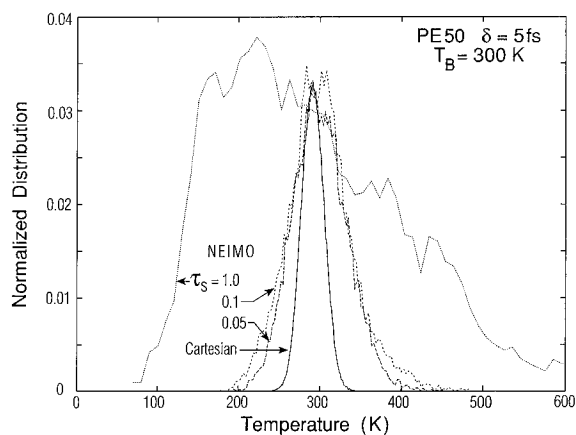
where each  $\text{CH}_3$  and  $\text{CH}_2$  group was taken as a cluster. This leads to 99 hinges and  $\mathcal{N} = 99 + 6 = 105$ , for the total number of degrees of freedom in the NEIMO–Hoover simulations (including the 6 degrees of freedom of the base body). Since the linear and angular momenta of the molecule are conserved, the temperature and other properties correspond to 99 independent dof.

Figure 2 shows the vibrational frequencies obtained from the velocity autocorrelation analysis of the Cartesian–Hoover dynamics for pe50. The modes below 200  $\text{cm}^{-1}$  (the first plateau shown in Figure 2 as a lattice) correspond to lattice vibrations. The region from around 800 to 1200  $\text{cm}^{-1}$  corresponds to  $\text{CH}_2$  rock twist and skeletal motions, as shown by Karasawa et al.<sup>22</sup> Also the region from 1300 to 1800  $\text{cm}^{-1}$  is due to  $\text{CH}_2$  wag- and scissor-like motion. The 3000  $\text{cm}^{-1}$  region is CH stretch. Thus the region between 200 and 600  $\text{cm}^{-1}$  corresponds to torsions. This leads to periods of 166–55 fs, suggesting a dynamics time step,  $\delta$ , of about 16–6 fs.

**5.1.1. Total Energy Conservation.** NEIMO–Hoover ( $T, V, \mathcal{N}$ ) simulations were carried out<sup>23</sup> for 400 ps at 300 K for various values of  $\tau_s$  and step size  $\delta$ . For  $\delta = 0.01$  ps, (38) suggests that the lower limiting value of  $\tau_s > 0.016$  ps, and indeed the NEIMO–Hoover simulations blew up for  $\tau_s = 0.007$  ps. Similarly for  $\delta = 0.005$  ps, (38) suggests that  $\tau_s > 0.008$  ps, and the simulations blew up for  $\tau_s = 0.007$  ps. On the basis of (41) we expect that for a total simulation time  $t_{\text{total}} = 400$  ps  $\tau_s$  should be  $\tau_s < 0.32$  ps. Indeed we found that the simulation



**Figure 3.** Variation with  $\tau_s$  of the rms fluctuations in the total energy (system + bath) for isolated pe50 as a function of  $\tau_s$ . Various integration time steps,  $\delta$ , were considered.



**Figure 4.** Distribution of temperatures  $\langle T_{\text{calc}} \rangle$  for pe50 from 400 ps of dynamics at  $T = 300$  K; Cartesian–Hoover simulations and NEIMO–Hoover simulations for various values of  $\tau_s$ .

with a large value of  $\tau_s = 10.0$  ps blew up. This was due to the slow response of the bath (large  $Q$ ) to the system.

For the fastest reliable results, we want to use the largest time step,  $\delta$ , consistent with stable dynamics. Of course a  $\delta$  that is too large can increase integration errors causing changes in the total energy and leading to errors in the calculated properties. From here onward total energy means the energy of the system and the bath. The root-mean-square (rms) fluctuations in the total energy are shown in Figure 3 for time steps of  $\delta = 1, 5,$  and  $10$  fs and various values of  $\tau_s$  (ranging from 0.01 to 1.0 ps). (Analysis of the rms fluctuations for larger time steps is given in the next section). For  $\delta = 1$  fs the total energy is well conserved at all values of  $\tau_s$ , apparently because  $\delta \ll \tau_s$ . For  $\delta = 5$  fs the total energy conservation is good for all values of  $\tau_s$  except 0.01. For  $\delta = 10$  fs the deviation in total energy is large only for  $\tau_s < 0.05$ . Thus,  $\tau_s = 0.05$  ps leads to good total energy conservation for all  $\delta \leq 10$  fs.

**5.1.2. Temperature Distributions.** We now compare the temperature distributions of the NEIMO–Hoover simulations with the temperature distribution for the Cartesian–Hoover simulation. Figure 4 shows the temperature distribution for 300 K Cartesian–Hoover simulations with  $\tau_s = 0.01$  ps,  $\delta = 1$  fs, and  $t_{\text{total}} = 400$  ps. Figure 4 also shows the temperature distribution for NEIMO–Hoover dynamics using various values of  $\tau_s$ . The NEIMO temperature distributions are similar to the Cartesian simulations for all  $\tau_s$ . For larger values of  $\tau_s$  the distribution is somewhat too broad. From statistical mechanics the average value of KE should be

$$\langle \text{KE} \rangle = \frac{N}{2} k T_B$$

and the fluctuation in KE should be

$$\langle (\delta \text{KE})^2 \rangle = \langle (\text{KE})^2 \rangle - \langle \text{KE} \rangle^2 = \frac{N}{2} (k T_B)^2$$

Hence the average temperature  $\langle T_{\text{calc}} \rangle$  should be

$$\langle T_{\text{calc}} \rangle = \frac{2}{Nk} \langle \text{KE} \rangle = T_B \quad (43)$$

and the mean-square deviation in temperature should be

$$\langle \delta T_{\text{calc}}^2 \rangle = \frac{2}{N} T_B^2$$

Thus we expect

$$\sqrt{\frac{N}{2} \langle \delta T_{\text{calc}}^2 \rangle} = T_B \quad (44)$$

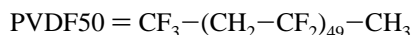
Figure 5a shows how fast  $\langle T_{\text{calc}} \rangle$  converges to the equilibrium value. The average temperature converges fast for  $\tau_s < 0.1$  ps but not for  $\tau_s = 0.5$  and 1.0. For  $\tau_s = 0.5$  and 1.0 there are large fluctuations even after 300 ps (they should dampen with longer simulation times). It is also evident from Figure 5a that  $\langle T_{\text{calc}} \rangle$  is steady only after 300 ps, even for  $\tau_s = 0.05$  ps. Hence the properties calculated below are averaged over the last 100 ps of the simulation. Table 1 shows  $\langle T_{\text{calc}} \rangle$  and  $((N/2) \delta T_{\text{calc}}^2)^{1/2}$  values for  $\delta = 5$  fs and various  $\tau_s$  values. According to (43) and (44) both  $\langle T_{\text{calc}} \rangle$  and  $((N/2) \delta T_{\text{calc}}^2)^{1/2}$  should be 300 K. We find that  $\langle T_{\text{calc}} \rangle$  is always slightly above 300 K but is closest for  $\tau_s = 0.05$ . The value of  $\langle \delta T_{\text{calc}}^2 \rangle$  is closest to the theoretical value for  $\tau_s = 0.07$  and 0.3 ps. However for  $\tau_s > 0.3$  ps, the fluctuations become increasingly too large. (38) for  $\delta = 5$  fs suggests that  $\tau_s > 0.008$  ps. These results suggest that  $\tau_s$  be chosen in the range of

$$0.05 \text{ ps} \leq \tau_s \leq 0.10 \text{ ps} \quad (45)$$

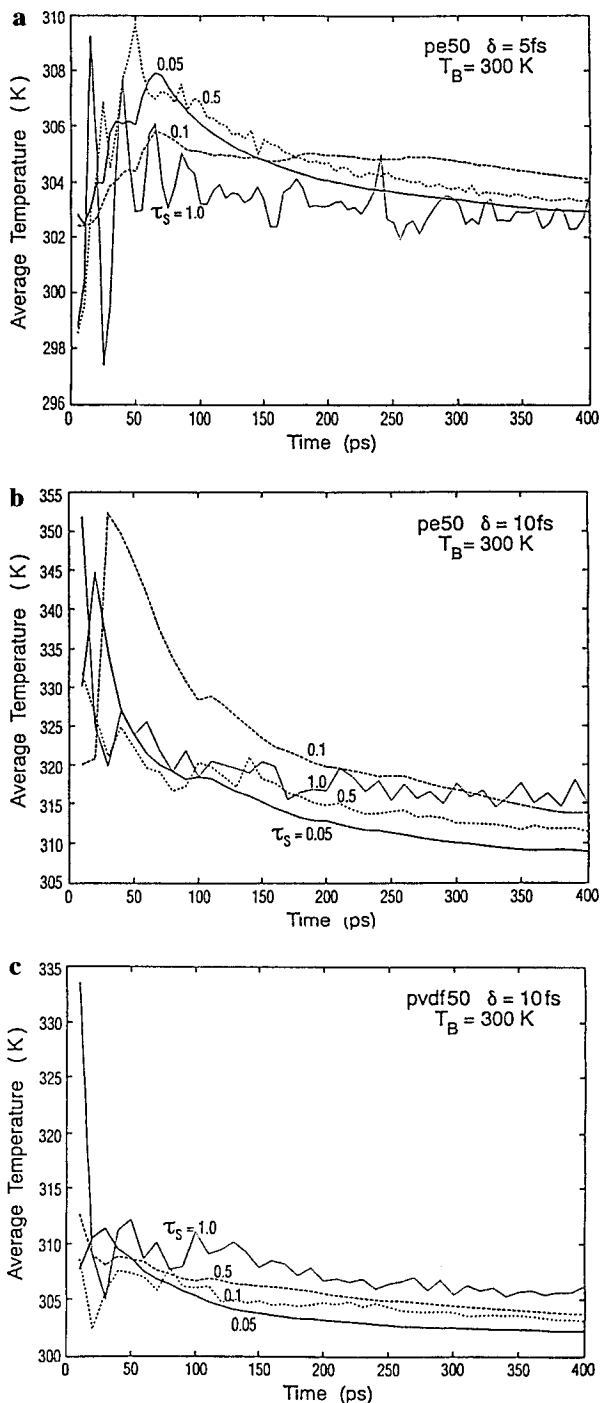
Further tests were carried out on pe50 with  $\delta = 10$  fs. (38) leads to  $\tau_s \geq 0.016$ . Indeed the simulations with  $\tau_s = 0.01$  ps blew up for this case. Table 2 shows the  $\langle T_{\text{calc}} \rangle$  and  $((N/2) \delta T_{\text{calc}}^2)^{1/2}$  where we see that  $\langle T_{\text{calc}} \rangle$  is too large by 3–6 K. This is because of inaccuracies in integration due to the large  $\delta$ . Once again good properties result from  $0.01 \text{ ps} < \tau_s < 0.1$  ps. Figure 5b shows the variation of  $\langle T_{\text{calc}} \rangle$  with time. The  $\langle T_{\text{calc}} \rangle$  converges fast for  $\tau_s < 0.1$  ps but for  $\tau_s = 1.0$  ps leads to large fluctuations even after 200 ps. Again the best values of  $\langle \delta T_{\text{calc}}^2 \rangle$  are obtained for

$$0.05 \text{ ps} \leq \tau_s \leq 0.07 \text{ ps} \quad (46)$$

## 5.2. PVDF-50. We now consider the polymer



to test the NEIMO–Hoover dynamics. This molecule was built with the POLYGRAF BUILDER, and we used the MSXX FF derived for amorphous poly(vinylidene fluoride) by Karasawa



**Figure 5.** Variation of  $\langle T_{\text{calc}} \rangle$  with time for NEIMO–Hoover simulations at 300 K at several values of  $\tau_s$ : (a) pe50 with  $\delta = 5$  fs, (b) pe50 with  $\delta = 10$  fs, and (c) PVDF50 with  $\delta = 10$  fs.

et al.<sup>25</sup> Each  $\text{CF}_3$ ,  $\text{CH}_3$ ,  $\text{CH}_2$ , and  $\text{CF}_2$  moiety was taken as a cluster connected by torsion-only hinges. This leads to 99 torsional dof. NEIMO–Hoover simulations were carried out for 400 ps with  $\delta = 10$  fs and various values of  $\tau_s$ . (38) demands that  $\tau_s > 0.016$ . As expected NEIMO–Hoover simulations for PVDF50 at 300 K blew up with  $\tau_s = 0.01$ .

Table 3 shows the  $\langle T_{\text{calc}} \rangle$  and  $((N/2) \delta T_{\text{calc}}^2)^{1/2}$  calculated for PVDF50. The  $\langle T_{\text{calc}} \rangle$  is close to  $T_B$  for all the values of  $\tau_s$  except for  $\tau_s = 10.0$  ps. Poor results were expected for  $\tau_s = 10.0$  ps since (41) suggests that  $t_{\text{total}} > 12600$  ps for good averaging. Both  $\langle T_{\text{calc}} \rangle$  and  $((N/2) \delta T_{\text{calc}}^2)^{1/2}$  are closest to 300 K for  $\tau_s = 0.1$  ps. For  $\tau_s > 0.3$  ps we see that  $((N/2) \delta T_{\text{calc}}^2)^{1/2}$  deviates increasingly from 300 K. Figure 5c shows the variation of  $\langle T_{\text{calc}} \rangle$  with simulation time for PVDF50 at 300

**TABLE 1: Dependence of Temperature Averages for Isolated pe50 as a Function of  $\tau_s$  with  $\delta = 5$  fs**

$\tau_s$ (ps)	$\langle T_{\text{calc}} \rangle$ (K)	$((N/2)\langle \delta T_{\text{calc}}^2 \rangle)^{1/2}$ (K)
0.007	blew up	
0.01	303.16	240.15
0.03	302.13	156.43
0.05	301.68	238.96
0.07	302.74	295.31
0.1	302.32	285.88
0.3	302.95	304.35
0.5	302.13	367.47
0.8	302.46	354.77
1.0	304.96	923.79
10.0	blew up	

**TABLE 2: Dependence of Temperature Averages for Isolated pe50 as a Function of  $\tau_s$  with  $\delta = 10$  fs**

$\tau_s$ (ps)	$\langle T_{\text{calc}} \rangle$ (K)	$((N/2)\langle \delta T_{\text{calc}}^2 \rangle)^{1/2}$ (K)
0.01	blew up	
0.03	303.93	142.39
0.05	305.71	223.08
0.07	305.29	276.19
0.1	304.99	335.41
0.3	303.55	705.89
0.5	308.10	1591.40
0.8	307.58	1868.88
1.0	306.93	2829.92
10.0	blew up	

**TABLE 3: Dependence of Temperature Averages for Isolated PVDF50 as a Function of  $\tau_s$  with  $\delta = 10$  fs**

$\tau_s$ (ps)	$\langle T_{\text{calc}} \rangle$ (K)	$((N/2)\langle \delta T_{\text{calc}}^2 \rangle)^{1/2}$ (K)
0.01	blew up	
0.03	301.67	120.91
0.05	301.54	311.23
0.07	302.10	181.99
0.1	301.66	160.00
0.3	302.45	389.79
0.5	301.02	548.20
0.8	302.84	662.94
1.0	308.60	1134.51
10.0	379.21	28501.37

K. The fluctuations in the average temperature converges faster for  $\tau_s = 0.05$  ps than for  $\tau_s = 1.0$  ps. As expected from (41) the fluctuations for  $\tau_s = 1.0$  ps are large even after 250 ps. Again, the best properties are obtained for

$$0.05 < \tau_s < 0.10 \quad (47)$$

**5.3. Conclusion.** On the basis of (45), (46), and (47) we recommend that

$$0.05 < \tau_s < 0.07 \quad (48)$$

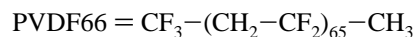
## 6.0. Applications of NEIMO–Hoover Dynamics

In this section we test the NEIMO–Hoover simulations for various systems using  $\tau_s = 0.05$  ps from (45). However for  $\delta = 25$  or 30 fs, we have used  $\tau_s = 0.07$  ps. From (38) this is consistent for  $\delta < 44$  fs. We now examine the limits on  $\delta$  for NEIMO–Hoover simulations of dense amorphous polymers (simulated using periodic boundary conditions).

**6.1. Isolated Polymer Chains.** NEIMO–Hoover simulations were carried out for isolated polyethylene oligomers: pe20, pe30, pe40, and pe50. Simulations<sup>11</sup> were done at 300 K for 400 ps. Again, only torsional dof were allowed (bond lengths and bond angles frozen). The total number of dof (including

base body) in the NEIMO–Hoover simulations for pe20, pe30, pe40, and pe50 are 45, 65, 85, and 105, respectively. Cartesian–Hoover simulations<sup>11</sup> were also done for all these systems at 300 K. Figure 6a shows how the time step  $\delta$  affects the total energy (scaled by the torsional dof). For NEIMO–Hoover dynamics, time steps as high as 25–30 fs give stable dynamics while Cartesian–Hoover simulations are limited to  $\delta \leq 3$  fs. Parts b, c, and d of Figure 6 show similar curves for pe30, pe40, and pe50, respectively. Again, time steps of 25–30 fs can be used for these systems, whereas Cartesian–Hoover simulations are limited to time steps  $\leq 2$  fs. Thus with NEIMO–Hoover the time steps can be about 10 times as large as for Cartesian dynamics while leading to the same total energy conservation.

**6.2. Amorphous Polymers.** Using periodic boundary conditions to eliminate surface effects, we examined NEIMO–Hoover dynamics for amorphous polymers poly(vinylidene fluoride)



with one chain per unit cell and poly(vinyl chloride)



with four chains per unit cell.

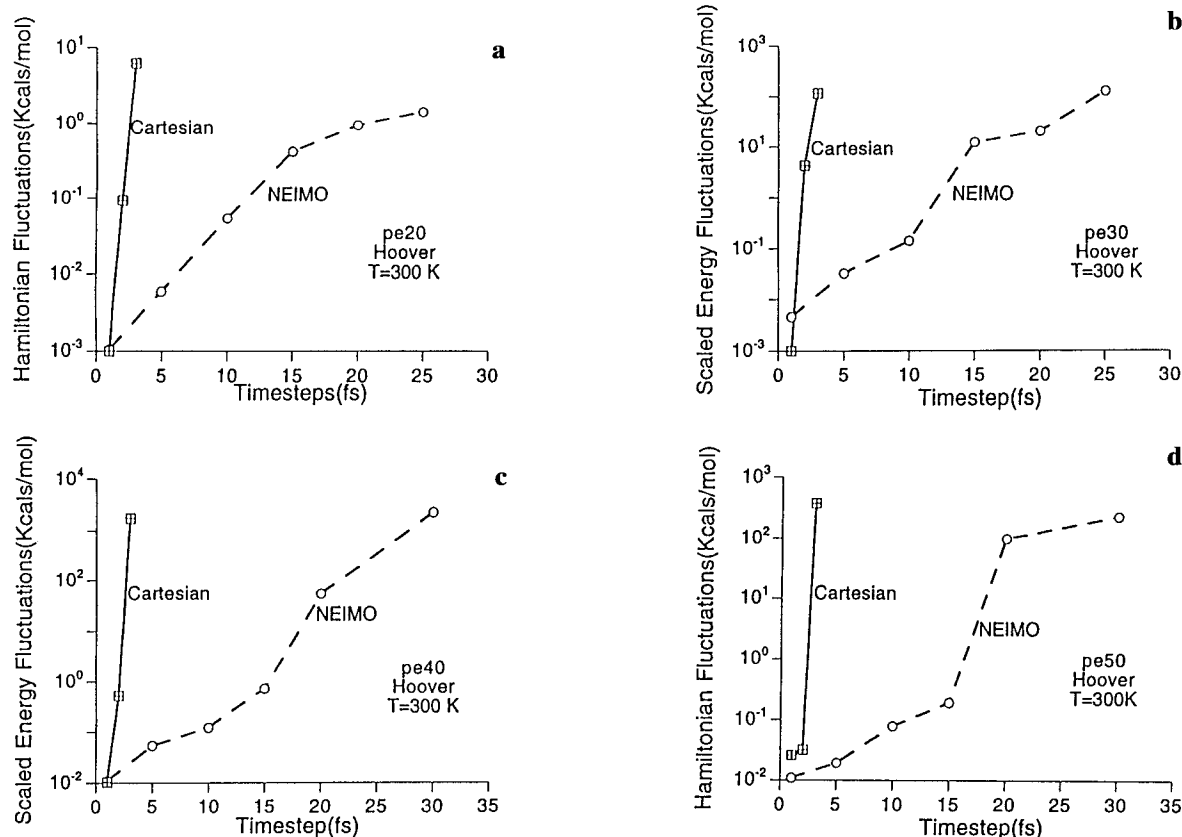
**6.2.1. PVDF66-PBC.** For amorphous PVDF66 (132 carbons) the unit cell is cubic with  $a = b = c = 18.0$  Å. This system has 131 torsional dof and a total of  $N = 137$ . Each backbone carbon with its ligands is taken as a cluster. We used the FF of Karasawa et al.<sup>25</sup> The nonbond forces were calculated using the fast Ewald summation method.<sup>26</sup> The initial structure was energy minimized (conjugate gradient) to a rms force less than 0.1 (kcal/mol)/Å. NEIMO–Hoover simulations were carried out at 300 K using  $\tau_s \geq 0.05$  ps. According to (38),  $\tau_s \geq 0.048$  ps for  $\delta = 30$  fs. The Cartesian–Hoover simulations were carried out for amorphous PVDF at the same conditions as those for NEIMO–Hoover. Figure 7 shows the total energy fluctuations for PVDF66 using various  $\delta$  values. We see that the  $\delta$  with NEIMO–Hoover can be 10 times as fast as that for Cartesian–Hoover with the same fluctuation in the total energy.

**6.2.2. (PVC20)<sub>4</sub>PBC.** Both Cartesian and NEIMO–Hoover simulations were carried out for amorphous PVC at 300 K using the Ewald summation method for nonbonds. Once again  $\tau_s = 0.05$  ps for all  $\delta$ . For  $\delta = 30$  fs,  $\tau_s = 0.07$  ps was used. For PVC the unit cell<sup>11</sup> has  $a = 22.8$  Å,  $b = 23.1$  Å,  $c = 12.1$  Å, and  $\alpha = 94.82^\circ$ ,  $\beta = 89.17^\circ$ , and  $\gamma = 84.92^\circ$ . The system has four chains of PVC with a total of 488 atoms and 180 torsional dof including the four base bodies for each chain. In order to remove any bad contacts in the starting structure, we minimized the energy of the initial structure to a rms in force less than 0.1 (kcal/mol)/Å. Figure 8 shows the fluctuations in the total energy as a function of time step for amorphous PVC. Again the time steps for NEIMO can be about 10 times those for Cartesian dynamics. Here we show that  $\delta = 20$  fs exhibits stable dynamics.

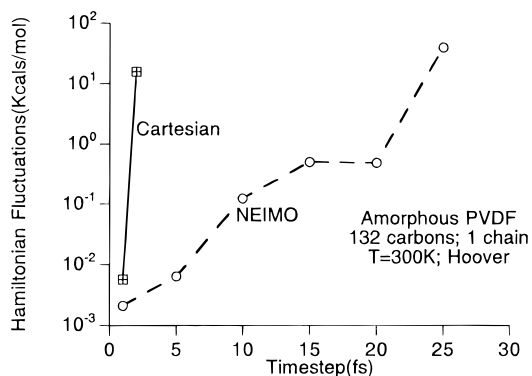
## 7.0. Summary

For polymers and proteins<sup>4</sup> NEIMO allows an order of magnitude speedup in the calculations since much larger integration time steps can be used than for Cartesian simulations. The NEIMO method has been extended to sample the  $T, V, N$  ensemble in order to obtain canonical ensembles for calculating properties of materials. The Lagrangian equations of motion have been derived for both the NEIMO–Nosé and the NEIMO–Hoover methods, leading to an additional frictional force in the Coriolis term. We have shown the applicability of the  $\mathcal{Q}\mathcal{N}$

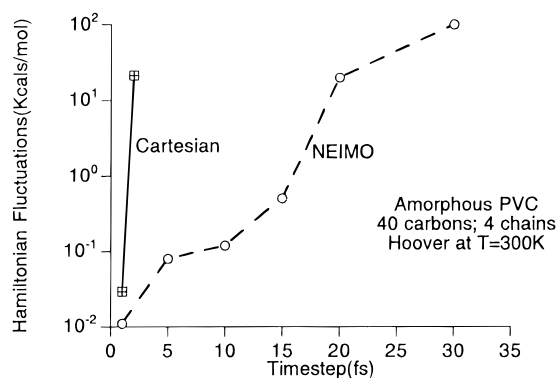




**Figure 6.** Fluctuations in total energy scaled by the number of torsional degrees of freedom for isolated PE chains at 300 K NEIMO–Hoover dynamics with various  $\delta$ : (a) pe20, (b) pe30, (c) pe40, and (d) pe50.



**Figure 7.** Fluctuations in total energy scaled by the number of torsional degrees of freedom at 300 K NEIMO–Hoover dynamics of PVDF-66 using  $\tau_s = 0.05$  ps and various  $\delta$  values. For  $\delta = 30$  fs, we used  $\tau_s = 0.07$  ps.



**Figure 8.** Fluctuations in total energy scaled by the number of torsional degrees of freedom at 300 K NEIMO–Hoover dynamics (PVC-20)<sub>4</sub> using  $\tau_s = 0.05$  ps and various  $\delta$  values. For  $\delta = 30$  fs, we used  $\tau_s = 0.07$  ps.

algorithm for solving these equations of motion. The significance of the choice of the heat bath relaxation constant  $\tau_s$  (i.e., the mass  $Q$ ) for doing torsional constrained molecular dynamics using the NEIMO algorithm has been established by performing NEIMO–Hoover simulations on pe50 and PVDF50 with various time steps. It has been shown that time steps as large as 20–30 fs can be used for NEIMO–Hoover simulations of amorphous polymers.

**Acknowledgment.** The research was funded by NSF (CHE 94-13930 and ACS 92-17368). The facilities of the MSC are also supported by grants from DOE-BCTR, Asahi Chemical, Asahi Glass, BP Chemical, Chevron Petroleum Technology, BF Goodrich, Xerox, Hughes Research Lab., and Beckman Institute. This work has been partially performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract

with the National Aeronautics and Space Administration. Some of the computations were carried out at the JPL CRAY facility, NSF Pittsburgh Supercomputing Center, and NSF San Diego Supercomputing Center.

## References and Notes

- (1) Mazur, A.; Abagyan, R. *J. Biomol. Struct. Dyn.* **1989**, *6*, 815.
- (2) Jain, A.; Vaidehi, N.; Rodriguez, G. *J. Comput. Phys.* **1993**, *106*, 258.
- (3) Rice, L. M.; Brunger, A. T. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 277.
- (4) Mathiowetz, A.; Jain, A.; Karasawa, N.; Goddard, W. A., III. *Proteins: Struct., Funct., Genet.* **1994**, *20*, 227. See also: Mathiowetz, A. Ph.D. Thesis, Department of Chemistry, California Institute of Technology, 1992.
- (5) Ding, H. Q.; Karasawa, N.; Goddard, W. A., III. *J. Chem. Phys.* **1992**, *97*, 4309.
- (6) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

- (7) Hoover, W. G. *Phys. Rev.* **1985**, *A31*, 1695.
- (8) Hoover, W. G. *Computational Statistical Mechanics*; Elsevier: Amsterdam, 1991.
- (9) Rodriguez, G.; Jain, A.; Kreutz-Delgado, K. *J. Astronaut. Sci.* **1992**, *40*, 27.
- (10) Rodriguez, G. *IEEE J. Rob. Automat.* **1987**, *3*, 624.
- (11) The NEIMO-Hoover calculations were carried out using POLYGRAF (from Materials Simulations Incorporated, Burlington, MA) modified to include the NEIMO option. All the polyethylene and amorphous PVC simulations were done with DREIDING FF.
- (12) Nosé, S. *Prog. Theor. Phys. Suppl.* **1991**, *103*, 1.
- (13) Ashurst, W. T.; Hoover, W. G. *Phys. Rev. Lett.* **1973**, *31*, 206.
- (14) Woodcock, L. V. *Chem. Phys. Lett.* **1971**, *10*, 257.
- (15) Abraham, F. F.; Koch, S. W.; Desai, R. C. *Phys. Rev. Lett.* **1983**, *78*, 2626.
- (16) Haile, S.; Gupta, S. *J. Chem. Phys.* **1983**, *79*, 3067.
- (17) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384.
- (18) Nosé, S. *Mol. Phys.* **1984**, *52*, 255.
- (19) Cagin, T.; Goddard, W. A., III; Ary, M. L. *Comput. Polym. Sci.* **1991**, *1*, 241.
- (20) DiTolla, F. D.; Ronchetti, M. *Phys. Rev. E* **1993**, *48*, 1726.
- (21) Cho, K.; Joannopoulos, J. D. *Phys. Rev. A* **1992**, *45*, 7089.
- (22) Karasawa, N.; Dasgupta, S.; Goddard, W. A., III. *J. Phys. Chem.* **1991**, *95*, 2260.
- (23) The pe50 molecule was built (using the BUILDER in POLYGRAF) to have a starting conformation with all the dihedrals in the trans conformation. The energy was minimized to a rms in force of 0.09 (kcal/mol/Å using DREIDING<sup>24</sup> FF and conjugate gradient minimization. The charge on H is 0.1440. A spline cutoff radius of 9 Å was used to calculate the nonbond forces with the inner and outer match radii being 8 and 8.5 Å, respectively. The nonbond pair list was updated every 0.1 ps of the simulation.
- (24) Mayo, S. I.; Olafson, B. D.; Goddard, W. A., III. *J. Phys. Chem.* **1990**, *94*, 8897.
- (25) Karasawa, N.; Goddard, W. A., III. *Macromolecules* **1992**, *25*, 7268.
- (26) Karasawa, N.; Goddard, W. A., III. *J. Phys. Chem.* **1989**, *93*, 7320.

JP9530430