

A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal–isobaric ensemble

Mark E Tuckerman^{1,2}, José Alejandro³, Roberto López-Rendón³,
Andrea L Jochim⁴ and Glenn J Martyna⁵

¹ Institute for Pure and Applied Mathematics (IPAM), University of California, Los Angeles, CA 90095, USA

² Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, New York, NY 10003, USA

³ Departamento de Química, Universidad Autónoma Metropolitana-Iztapalapa, Av San Rafael Atlixco 186, 09340, México DF, Mexico

⁴ Department of Chemistry, New York University, New York, NY 10003, USA

⁵ Physical Sciences Division, IBM TJ Watson Research Center, Yorktown Heights, NY 10598, USA

Received 24 October 2005, in final form 16 January 2006

Published 24 April 2006

Online at stacks.iop.org/JPhysA/39/5629

Abstract

The constant-pressure, constant-temperature (*NPT*) molecular dynamics approach is re-examined from the viewpoint of deriving a new measure-preserving reversible geometric integrator for the equations of motion. The underlying concepts of non-Hamiltonian phase-space analysis, measure-preserving integrators and the symplectic property for Hamiltonian systems are briefly reviewed. In addition, current measure-preserving schemes for the constant-volume, constant-temperature ensemble are also reviewed. A new geometric integrator for the *NPT* method is presented, is shown to preserve the correct phase-space volume element and is demonstrated to perform well in realistic examples. Finally, a multiple time-step version of the integrator is presented for treating systems with motion on several time scales.

PACS numbers: 05.20.–y, 02.40.–k, 31.15.Qg

1. Introduction

Over the last several decades, molecular dynamics (MD) has become one of the most important and commonly used approaches for studying condensed phase systems. In an MD simulation, it is often necessary or desirable to study a system under conditions of constant pressure P and temperature T . Together with constant particle number N , these conditions define the isothermal–isobaric or *NPT* ensemble. The *NPT* ensemble is often used for purposes

of equilibrating a system, allowing it to adjust to an appropriate density, for computing equilibrium properties under isobaric conditions, such as the Gibbs free energy, and for studying structural phase transitions.

Among the ensembles typically used for molecular dynamics, namely the microcanonical (*NVE*), canonical (*NVT*) and isothermal–isobaric (*NPT*) ensembles, the *NPT* ensemble is generally regarded as the most difficult to generate due to the requirement that both the kinetic energy and instantaneous pressure must fluctuate according to the ensemble distribution function, which means that both the total energy and volume should vary such that the ensemble distribution function

$$f(x, V) = \frac{1}{N!V_0h^{3N}} \frac{\exp[-\beta(H(x) + PV)]}{\Delta(N, P, T)} \quad (1.1)$$

is generated. Here, P , V and T are the applied external pressure, the system volume and the external temperature, respectively, $\beta = 1/kT$, x is the phase-space vector, i.e. the vector whose components are the momenta and coordinates of all of the particles in the system, and $H(x)$ is the system Hamiltonian. V_0 is an arbitrary reference volume, and h is Planck's constant. The quantity $\Delta(N, P, T)$ is the partition function of the ensemble given by

$$\Delta(N, P, T) = \frac{1}{N!V_0h^{3N}} \int_0^\infty dV \int dx \exp[-\beta(H(x) + PV)]. \quad (1.2)$$

If the Cartesian momenta and coordinates of the N particles are denoted by $\mathbf{p}_1, \dots, \mathbf{p}_N$, $\mathbf{r}_1, \dots, \mathbf{r}_N \equiv \mathbf{p}, \mathbf{r}$, respectively, then equation (1.2) becomes

$$\Delta(N, P, T) = \frac{1}{N!V_0h^{3N}} \int_0^\infty dV \int_{D(V)} d^N \mathbf{r} \int d^N \mathbf{p} [-\beta(H(\mathbf{p}, \mathbf{r}) + PV)] \quad (1.3)$$

where each of the coordinate integrations is restricted by the spatial domain defined by the volume at each value of the volume integration.

The most commonly used algorithms for generating the *NPT* ensemble are based on the ‘extended-phase-space’ approach [1, 9, 10, 17, 31], in which the volume is introduced as an additional dynamical variable, along with a corresponding momentum variable, in order to maintain isobaric conditions, while additional thermostat variables [21, 9, 18, 3, 15, 14] are used to control the kinetic energy fluctuations. Additional mass-like parameters are assigned to each of these additional variables, and it is these parameters which determine the time scales on which the latter evolve. Equations of motion are specified as a means of generating the phase-space distribution of the ensemble. Given that simulations in the *NPT* ensemble are typically more challenging than in the *NVE* or *NVT* ensembles, it is important to have robust, measure-preserving numerical methods for integrating the equations of motion. In this paper, such an algorithm is introduced.

Of the various dynamical schemes that have been proposed for generating the *NPT* ensemble, we will focus our attention on the algorithm of Martyna, Tobias and Klein (MTK) [17], which has been shown to correctly reproduce the isothermal–isobaric distribution of equation (1.1) [38]. We point out, however, that Hamiltonian formulations of the *NPT* ensemble have also been presented [31], although these will not be discussed here. We will begin the discussion by reviewing briefly the formalism underlying the statistical analysis of non-Hamiltonian dynamical systems [22, 41, 42, 38, 29, 26, 27, 34, 33, 28]. We will then review the algorithm of [19] for integrating the Nosé–Hoover chain (NHC) equations of motion for the *NVT* ensemble [18], as this algorithm will figure prominently in the new *NPT* method. Following this, we shall present the new approach for the MTK equations of motion [17]. We will discuss how to combine the new integration algorithm with the multiple time-step method

r-RESPA [40] for handling systems with multiple time-scale motion. Finally, a selection of examples will be presented illustrating the performance of the method on realistic problems.

2. Statistical analysis of non-Hamiltonian dynamical systems

The dynamical systems to be considered in this paper possess two important characteristics: (1) they involve ‘extended’ phase spaces in which the physical coordinates and momenta are supplemented by additional variables which regulate the fluctuations in the estimators that determine the thermodynamic control variables of a given statistical ensemble; and (2) they are non-Hamiltonian (By ‘non-Hamiltonian’, we generally mean truly non-Hamiltonian systems or Hamiltonian systems expressed in non-canonical coordinates.). A classical statistical mechanical theory for non-Hamiltonian systems was developed recently [42, 38, 29, 26, 27, 34, 33, 28]. This theory provides a framework for analysing non-Hamiltonian systems and determining the precise phase-space distribution generated by the equations of motion assuming ergodicity. The main points of this theory are reviewed below.

Consider a general dynamical system described by the equations of motion

$$\dot{x} = \xi(x) \quad (2.1)$$

where x is the complete phase-space vector and $\xi(x)$ is a vector field on the phase space. A sufficient condition to determine whether the equations of motion are non-Hamiltonian is that the phase-space compressibility

$$\kappa(x) = \nabla \cdot x = \nabla \cdot \xi(x) \quad (2.2)$$

be nonzero. This condition, however, is not necessary for a system to be non-Hamiltonian, there are non-Hamiltonian systems with vanishing compressibility. A nonvanishing compressibility has important implications for the structure of the phase space. In order to see this, consider a solution of equations (2.1) starting from an initial phase-space point x_0 . Denote this solution at time t by $x_t(x_0)$. The notation $x_t(x_0)$ indicates that the solution is a unique function of the initial conditions. This solution can be viewed as a transformation from phase-space coordinates at x_0 at $t = 0$ to new phase-space coordinates x_t at time t generated by the single parameter t . Thus, the transformation can also be regarded as a one-parameter family of diffeomorphisms on the phase space. Let $J(x_t; x_0) = |\partial x_t / \partial x_0|$ be the Jacobian of this transformation. The Jacobian determines how the initial phase-space volume measure dx_0 transforms to dx_t :

$$dx_t = J(x_t; x_0) dx_0. \quad (2.3)$$

It is straightforward to show that the Jacobian satisfies an equation of motion of the form [2, 42]

$$\frac{d}{dt} J(x_t; x_0) = \kappa(x_t) J(x_t; x_0) \quad (2.4)$$

subject to an obvious initial condition $J(x_0; x_0) = 1$. Equation (2.4) can be solved by the method of characteristics to yield

$$J(x_t; x_0) = \exp \left[\int_0^t ds \kappa(x_s) \right]. \quad (2.5)$$

If equation (2.1) are incompressible, as would be the case for a Hamiltonian system, then equation (2.5) implies that $J(x_t; x_0) = 1$ for all time, leading to Liouville’s theorem $dx_t = dx_0$, which states that the phase-space volume measure is conserved. On the other hand, if $\kappa(x) \neq 0$, then the Jacobian will not be unity. However, equation (2.4) implies that $\kappa(x_t) = d \ln J(x_t; x_0) / dt$. Of course, substituting this into equation (2.5) leads to a tautology,

but if we let $w(x_t, t)$ be a function such that $\kappa(x_t) = dw(x_t, t)/dt$, then equation (2.5) can be expressed as

$$J(x_t; x_0) = e^{w(x_t, t) - w(x_0, 0)} = \frac{e^{-w(x_0, 0)}}{e^{-w(x_t, t)}}. \quad (2.6)$$

Now, if the phase space is allowed to have a nontrivial metric G_{ij} with the determinant $g = \det(G)$, then the Jacobian can be expressed in the familiar manner as a ratio of the square root of metric determinant factors [6, 7, 25]

$$J(x_t; x_0) = \frac{\sqrt{g(x_0, 0)}}{\sqrt{g(x_t, t)}} \quad (2.7)$$

and we can identify $\sqrt{g(x_t, t)} = \exp(-w(x_t, t))$. Substitution of equation (2.7) into equation (2.3) leads to a volume conservation law on a phase space with a nontrivial metric:

$$\sqrt{g(x_t, t)} dx_t = \sqrt{g(x_0, 0)} dx_0 \quad e^{-w(x_t, t)} dx_t = e^{-w(x_0, 0)} dx_0. \quad (2.8)$$

Equation (2.7) implies that the metric determinant $\sqrt{g(x_t, t)}$ satisfies an equation of motion of the form

$$\frac{d}{dt} \sqrt{g(x_t, t)} = -\kappa(x_t) \sqrt{g(x_t, t)}. \quad (2.9)$$

Note that if the compressibility of the system is zero, then $J(x_t, x_0) = 1$, and $w(x_t, t) = 0$ and $\sqrt{g(x_t, t)} = 1$ for all time, and equation (2.8) reduces to the usual volume conservation law $dx_0 = dx_t$. This shows that a system need not be Hamiltonian for usual volume conservation law to apply. The preceding discussion allows for the most general situation, i.e. one in which the metric possesses an explicit time dependence. The role of explicitly time-dependent metrics in the analysis of dynamical systems has been discussed elsewhere [23, 24]. However, the systems that will be considered involve metric factors that only depend on the phase-space coordinates x with no explicit time dependence.

If the metric is properly included in the definition of the phase space, then it can be shown that there exists a phase-space probability density $f(x_t, t)$ corresponding to an ensemble of trajectories generated by equations (2.1) that is conserved, i.e. $df/dt = 0$ [42, 38]. If the ensemble is an equilibrium ensemble, then $\partial f/\partial t = 0$, so that $f = f(x)$ has no explicit time dependence. Moreover, if equations (2.1) possess N_c conservation laws of the form $\Lambda_k(x) = C_k$, $k = 1, \dots, N_c$, then, provided the motion described by equations (2.1) is ergodic, the ensemble distribution that will be generated is of a general ‘microcanonical’ form [38]:

$$f(x) = \mathcal{N} \prod_{k=1}^{N_c} \delta(\Lambda_k(x) - C_k) \quad (2.10)$$

where \mathcal{N} is a normalization constant. Thus, if the phase-space vector corresponds to an extended phase space, involving the physical system variables and additional unphysical variables, then when equation (2.10) is integrated over the latter, the phase-space distribution in the physical variables is generated by the equations of motion can be determined. This technique was used in [38] to analyse a wide variety of existing molecular dynamics algorithms in order to determine which existing methods actually produced the ensemble distributions they purported to with somewhat unexpected results which were also confirmed by numerical simulations.

As noted earlier, the existence of a metric determinant $\sqrt{g(x_t, t)}$ on the phase space implies the existence of a metric tensor $G_{ij}(x, t)$. Recently, Tarasov derived an equation of motion for the full metric tensor that is consistent with equation (2.9) [33]. Denoting the

components of x as $x = (x^1, \dots, x^n)$ in an n -dimensional phase space, and the components of the vector field $\xi(x)$ as $\xi(x) = (\xi^1(x), \dots, \xi^n(x))$ (the use of contravariant indices is intentional), the equation of motion for the phase-space metric tensor is [33]

$$\frac{\partial G_{ij}}{\partial t} = \sum_k \left[\frac{\partial}{\partial x^i} (G_{jk} \xi^k) - \frac{\partial}{\partial x^j} (G_{ik} \xi^k) \right] \tag{2.11}$$

which yields a skew-symmetric metric as required on the phase space. This equation of motion guarantees that the Jacobi identity

$$\frac{\partial}{\partial x^i} G_{jk}(x, t) + \frac{\partial}{\partial x^j} G_{ki}(x, t) + \frac{\partial}{\partial x^k} G_{ij}(x, t) = 0 \tag{2.12}$$

is satisfied. The importance of the phase-space metric tensor is its use in the definition of a Poisson bracket between two phase-space functions $A(x)$ and $B(x)$ on the non-Hamiltonian phase space, which is given by

$$\{A, B\} = \sum_{i,j} \frac{\partial A}{\partial x^i} G^{ij}(x, t) \frac{\partial B}{\partial x^j} \tag{2.13}$$

where $G^{ij}(x, t)$ and $G_{ij}(x, t)$ are related by

$$\sum_k G^{ik}(x, t) G_{kj}(x, t) = \delta_j^i. \tag{2.14}$$

A final note concerns the role of the metric in the symplectic condition. Consider the metric tensor M for a Hamiltonian system

$$M = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \tag{2.15}$$

where I is an identity matrix of dimension $n/2$. The symplectic condition states the following: if J denotes the Jacobian matrix of the transformation $x_0 \rightarrow x_t$, then

$$JMJ^T = M \tag{2.16}$$

where J^T is the transpose of the Jacobian matrix. Written out in terms of matrix elements, equation (2.16) becomes

$$M_{ij} = \frac{\partial x_t^k}{\partial x_0^i} \frac{\partial x_t^l}{\partial x_0^j} M_{kl}. \tag{2.17}$$

However, the right side of this equation determines how the metric tensor transforms under a coordinate transformation $x_0 \rightarrow x_t$. Thus, the symplectic condition is a statement that the time evolution of the phase space of a Hamiltonian system preserves the form of the metric tensor. It is, therefore, reasonable to assume that a properly defined metric tensor $G_{ij}(x)$ on a non-Hamiltonian phase space would preserve its form under the exact time evolution, i.e.

$$G_{ij}(x_0) = \frac{\partial x_t^k}{\partial x_0^i} \frac{\partial x_t^l}{\partial x_0^j} G_{kl}(x_t). \tag{2.18}$$

Equation (2.18) can be viewed as a generalization of the symplectic condition for a non-Hamiltonian phase space. A proof of this result for flat phase spaces in arbitrary coordinates was provided in [11]. The interested reader is referred to this work for detail of the proof. A more general proof of this result is beyond the scope of the present paper but will be presented in a forthcoming publication [36].

The above formalism defines the main objects needed to develop and analyse techniques in non-Hamiltonian phase spaces. Although the discussion in the subsequent sections will focus on geometric integration, we will need to refer to this formalism and its previously reported application [38].

3. Geometric integration and the Liouville operator

The approach that will be presented here for the development of geometric integrators of non-Hamiltonian systems will be based on a generalization of the Liouville operator. The use of this technique has been discussed extensively in the literature (see, for example, [40, 19]) and will be used here to develop a new approach for the equations of motion of the *NPT* ensemble. Before discussing this case, however, let us first review briefly the fundamental concept.

The equations of motion equation (2.1) can be written in the form of an operator equation

$$\dot{x} = iLx \quad (3.1)$$

where

$$iL = \xi(x) \cdot \nabla \quad (3.2)$$

is a generalization of the Liouville operator. When the equations of motion are expressed as in equation (3.1), the evolution of an arbitrary initial condition x_0 can be determined formally according to

$$x_t = e^{iLt} x_0 \quad (3.3)$$

where the operator $\exp(iLt)$ is known as the classical propagator. Although equation (3.3) is only a formal solution, it is, nevertheless, the starting point for the development of geometrical integration schemes.

Consider a Hamiltonian system expressed in terms of Cartesian coordinates $\mathbf{r}_1, \dots, \mathbf{r}_N \equiv \mathbf{r}$ and momenta $\mathbf{p}_1, \dots, \mathbf{p}_N \equiv \mathbf{p}$. If $H(\mathbf{p}, \mathbf{r})$ is the Hamiltonian, the equations of motion will be given by Hamilton's equations

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} \quad \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i}. \quad (3.4)$$

In this case, equation (3.2) becomes the true Liouville operator

$$iL = \sum_{i=1}^N \left[\frac{\partial H}{\partial \mathbf{p}_i} \cdot \frac{\partial}{\partial \mathbf{r}_i} - \frac{\partial H}{\partial \mathbf{r}_i} \cdot \frac{\partial}{\partial \mathbf{p}_i} \right]. \quad (3.5)$$

Thus, if $H(\mathbf{p}, \mathbf{r})$ is of the typical form

$$H(\mathbf{p}, \mathbf{r}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (3.6)$$

where $U(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is an N -body potential, then equation (3.5) becomes

$$iL = \sum_{i=1}^N \left[\frac{\mathbf{p}_i}{m_i} \cdot \frac{\partial}{\partial \mathbf{r}_i} + \mathbf{F}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \right] \quad (3.7)$$

where $\mathbf{F}_i = -\partial U / \partial \mathbf{r}_i$ is the force on particle i .

Although the propagator cannot be evaluated exactly, note that the operator in equation (3.7) naturally separates into two contributions, $iL = iL_1 + iL_2$, where

$$iL_1 = \sum_{i=1}^N \frac{\mathbf{p}_i}{m_i} \cdot \frac{\partial}{\partial \mathbf{r}_i} \quad iL_2 = \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial}{\partial \mathbf{p}_i}. \quad (3.8)$$

This separation allows the Trotter [35, 5] theorem to be applied

$$e^{iLt} = \lim_{M \rightarrow \infty} [e^{iL_2 t/2M} e^{iL_1 t/M} e^{iL_2 t/2M}]^M \quad (3.9)$$

where M is an integer. For finite M , the Trotter theorem leads to an approximate propagator for a small time interval Δt :

$$e^{iL\Delta t} = e^{iL_2\Delta t/2} e^{iL_1\Delta t} e^{iL_2\Delta t/2} + \mathcal{O}(\Delta t^3). \quad (3.10)$$

The symmetric form of this factorization ensures that the resulting integrator will be symplectic as required for a Hamiltonian system. Note that, although the error in one step is of order Δt^3 , the global error in a long trajectory is of order Δt^2 . It can be shown that the application of the factorized operator in equation (3.10) on the initial phase space yields the familiar velocity Verlet integrator [40, 19]:

$$\begin{aligned} \mathbf{r}_i(\Delta t) &= \mathbf{r}_i(0) + \Delta t \frac{\mathbf{p}_i(0)}{m_i} + \frac{\Delta t^2}{2m_i} \mathbf{F}_i(\mathbf{r}(0)) \\ \mathbf{p}_i(\Delta t) &= \mathbf{p}_i(0) + \frac{\Delta t}{2} [\mathbf{F}_i(\mathbf{r}(0)) + \mathbf{F}_i(\mathbf{r}(\Delta t))]. \end{aligned} \quad (3.11)$$

The only identity needed to derive this result is the action of an operator of the form $\exp(c\partial/\partial y)$, for c independent of y , on a function $f(y)$:

$$\exp\left(c \frac{\partial}{\partial y}\right) f(y) = f(y + c) \quad (3.12)$$

which demonstrates that the operator acts as a translation operator. By computing the Jacobian matrix of equations (3.11), it can be shown that this integrator is symplectic and, therefore, preserves the phase-space volume. In addition, it is not necessary to derive closed-form expressions for integrators based on factorizations of the classical propagator, such as equation (3.10). In fact, the same algorithm results if each operator in equation (3.10) is ‘translated’ directly into an instruction in computer code, in this case, each instruction is a simple translation either of position or momentum. This idea was introduced in [19] and is known as the *direct translation* technique.

We note, additionally, that the Liouville operator formalism allows systems with motion on several time scales to be integrated using a multiple time-step approach [40]. Suppose the force \mathbf{F}_i can be expressed as a sum of two components $\mathbf{F}_i^{(1)} + \mathbf{F}_i^{(2)}$, where $\mathbf{F}_i^{(1)}$ is a rapidly varying force and $\mathbf{F}_i^{(2)}$ is slowly varying. We seek to introduce two time steps δt and $\Delta t = n\delta t$ appropriate for the fast and slow components, respectively, into a geometric integration scheme. In order to accomplish this, we define the following contributions to the Liouville operator:

$$iL_1^{(1)} = \sum_{i=1}^N \frac{\mathbf{p}_i}{m_i} \cdot \frac{\partial}{\partial \mathbf{r}_i} \quad iL_1^{(2)} = \sum_{i=1}^N \mathbf{F}_i^{(1)} \cdot \frac{\partial}{\partial \mathbf{p}_i} \quad iL_2 = \sum_{i=1}^N \mathbf{F}_i^{(2)} \cdot \frac{\partial}{\partial \mathbf{p}_i} \quad (3.13)$$

and factorize the propagator according to the two time-step scheme:

$$e^{iL\Delta t} = e^{iL_2\Delta t/2} (e^{iL_1^{(2)}\delta t/2} e^{iL_1^{(1)}\delta t} e^{iL_1^{(2)}\delta t/2})^n e^{iL_2\Delta t/2} + \mathcal{O}(\Delta t^3) \quad (3.14)$$

which is known as the reversible reference system propagator algorithm (r-RESPA) [40]. As the name suggests, the two contributions to iL_1 act as a kind of ‘reference system’ described by the force $\mathbf{F}_i^{(1)}$. Thus, the operator in brackets is applied for n steps of size δt according to the scheme of equation (3.11) with the operator $\exp(iL_2\delta t/2)$ acting as a momentum correction factor before and after this reference system integration is performed. Since many problems in molecular dynamics involve computationally inexpensive fast forces and expensive slow forces, the use of such a scheme can reduce the overall computational overhead of a calculation by reducing the total number of slow force evaluations.

The theoretical underpinning of the above approach lies in the fact that the error can be proved to be bounded. Denoting the factorized propagator in equation (3.10) as $\exp(i\tilde{L}\Delta t)$, it can be shown from the Baker–Campbell–Hausdorff (BCH) formula [35, 43] that

$$\exp(i\tilde{L}\Delta t) = \exp\left[\Delta t\left(iL + \sum_{k=1}^{\infty}\Delta t^{2k}C_k\right)\right] \quad (3.15)$$

where the corrections C_k are expressible in terms of nested commutators of the Liouville operators iL_1 and iL_2 . For example, it can be shown that C_1 is given by

$$C_1 = \frac{1}{24}[iL_2 + 2iL_1, [iL_2, iL_1]]. \quad (3.16)$$

More importantly, it can be shown [4] that in a Hamiltonian system, commutators between Liouville operators yield new Liouville operators derivable from the Poisson brackets between the corresponding Hamiltonians. Thus, if $H = H_1 + H_2$ and $iL = iL_1 + iL_2$ are the corresponding contributions to the total Liouville operator, then $iL_3 = [iL_1, iL_2]$ is a Liouville operator derivable from the Hamiltonian $H_3 = \{H_1, H_2\}$. Thus, the operator in equation (3.16) is the Liouville operator for a Hamiltonian of the form

$$\tilde{H}_1 = \frac{1}{24}\{H_2 + 2H_1, \{H_2, H_1\}\}. \quad (3.17)$$

Similarly, every C_k in equation (3.15) is the Liouville operator corresponding to a Hamiltonian denoted \tilde{H}_k obtained from a nested Poisson bracket structure of the same form as the nested commutator structure that determines C_k . Consequently, there is a Hamiltonian $\tilde{H}(\mathbf{p}, \mathbf{r}; \Delta t)$ expressible as

$$\tilde{H}(\mathbf{p}, \mathbf{r}; \Delta t) = H(\mathbf{p}, \mathbf{r}) + \sum_{k=1}^{\infty}\Delta t^{2k}\tilde{H}_k(\mathbf{p}, \mathbf{r}) \quad (3.18)$$

that is *exactly* conserved by the propagator in equation (3.15). Note that this Hamiltonian depends on the time step Δt , reduces to the true Hamiltonian in the limit $\Delta t \rightarrow 0$ and only exists for symplectic propagation schemes. The dynamics of the propagator $\exp(i\tilde{L}t)$ is exactly specified by Hamilton's equations in the form

$$\dot{\mathbf{r}}_i = \frac{\partial \tilde{H}}{\partial \mathbf{p}_i} \quad \dot{\mathbf{p}}_i = -\frac{\partial \tilde{H}}{\partial \mathbf{r}_i}. \quad (3.19)$$

The existence of a Hamiltonian $\tilde{H}(\mathbf{p}, \mathbf{r}; \Delta t)$ means that the error in the integrator is bounded, since $|\tilde{H}(\mathbf{p}, \mathbf{r}; \Delta t) - H(\mathbf{p}, \mathbf{r})|$ is, itself, bounded, and $\tilde{H}(\mathbf{p}, \mathbf{r}; \Delta t)$ is exactly conserved. As an example, an exact closed-form expression for \tilde{H} exists for a simple harmonic oscillator of mass m and frequency ω . If p and q are the momentum and coordinate of the oscillator, and the Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2q^2 \quad (3.20)$$

then it can be shown that the Hamiltonian \tilde{H} that is exactly conserved by the velocity Verlet integrator is given by

$$\tilde{H}(p, q; \Delta t) = \left(\frac{p^2[1 - (\frac{\omega\Delta t}{2})^2]^{-1/2}}{2m} + \frac{m\omega^2q^2[1 - (\frac{\omega\Delta t}{2})^2]^{1/2}}{2}\right)\frac{\arccos(1 - \frac{\omega^2\Delta t^2}{2})}{|\omega\Delta t|}. \quad (3.21)$$

Although a closed form for \tilde{H} exists for the harmonic oscillator, convergence of equation (3.18) for nonlinear systems is generally not guaranteed [8], in which case, a truncated series can be employed in practice [30, 12, 8].

Before closing this section, we note that the above formalism has not been fully generalized for non-Hamiltonian systems, although presumably such a generalization could be accomplished if a procedure for determining the metric tensor of a non-Hamiltonian phase space existed. At present, no such a method exists. It is, nevertheless, clear that a properly formulated geometric integrator for a non-Hamiltonian system should, as a minimum condition, preserve the volume conservation law equation (2.8). It has been shown that an approach based on the generalized Liouville operator can be made to satisfy this condition while those derived from a Taylor series expansion generally cannot [39]. Indeed, various algorithms derived from a Taylor series approach have been introduced in the literature [18, 17, 13] and are currently in use, yet they do not satisfy equation (2.8). Nevertheless, it is also clear that additional fundamental work is needed to generalize the Liouville-operator based approach to arbitrary non-Hamiltonian systems in order to ensure that volume-preserving factorizations can be generated via a well-defined procedure, since *ad hoc* factorizations may well fail to satisfy equation (2.8). However, until such a theory is available, it is always necessary to check each integration scheme carefully for its adherence to equation (2.8). In what follows, the use of factorization schemes on the classical propagator will be employed, and appropriate checks will be presented.

4. Integrating the Nosé–Hoover chain equations for the NVT ensemble

In the molecular dynamics literature, a well-known dynamical scheme that purports to generate a canonical distribution is the so-called Nosé–Hoover method [9]. Although this method can be shown to generate a proper NVT ensemble under very special circumstances [9, 18, 38], the failure of this approach under the conditions usually used in molecular dynamics calculations was made clear in [38]. It was shown by Martyna *et al* that the shortcomings of the Nosé–Hoover algorithm could be overcome by a generalization of the approach in the form of the Nosé–Hoover chain equations [18]. In this scheme, the physical phase space is supplemented by a set of heat-bath variables $\eta_1, \dots, \eta_M, p_{\eta_1}, \dots, p_{\eta_M}$, which serve to drive the fluctuations of the kinetic energy in such a way that they average to the proper canonical value. The equations of motion are

$$\begin{aligned} \dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} & \dot{\mathbf{p}}_i &= \mathbf{F}_i - \frac{p_{\eta_1}}{Q_1} \mathbf{p}_i & \dot{\eta}_k &= \frac{p_{\eta_k}}{Q_k} \quad k = 1, \dots, M \\ \dot{p}_{\eta_k} &= G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} & \dot{p}_{\eta_M} &= G_M \end{aligned}$$

where the heat-bath ‘forces’ are given by

$$G_1 = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - 3NkT \quad G_k = \frac{p_{\eta_{k-1}}^2}{Q_{k-1}} - kT. \quad (4.1)$$

The parameters Q_1, \dots, Q_M are mass-like parameters (having units of energy \times time²) that determine the time scale on which the heat-bath variables evolve. The physics embodied in equations (4.1) is based on the fact that the term $-(p_{\eta_1}/Q_1)\mathbf{p}_i$ in the momentum equation acts as a kind of dynamic frictional force. Although the average $\langle p_{\eta_1} \rangle = 0$, instantaneously, p_{η_1} can be positive or negative and, therefore, act to damp or boost the momentum. According to the equation for p_{η_1} , if the kinetic energy is larger than $3NkT/2$, p_{η_1} will increase and have a greater damping effect on the momenta, while if the kinetic energy is less than $3NkT/2$, p_{η_1} will decrease and have a greater boosting effect on the momenta. In this way, the NHC system acts as a ‘thermostat’ regulating the kinetic energy so that its average is the correct canonical

value. In a similar manner, the $(k + 1)$ th heat-bath variable serves to modulate the fluctuations in the k th variable so that each heat-bath variable (except the M th variable) is ‘driven’ to have a proper canonical average.

Equations (4.1) have the conserved energy

$$H' = H(\mathbf{p}, \mathbf{r}) + \sum_{k=1}^M \frac{p_{\eta_k}^2}{2Q_k} + 3NkT\eta_1 + kT \sum_{k=2}^M \eta_k \quad (4.2)$$

and a compressibility

$$\kappa(x) = -3N\dot{\eta}_1 - \sum_{k=2}^M \dot{\eta}_k \quad (4.3)$$

and therefore a phase-space metric $\sqrt{g} = \exp(3N\eta_1 + \dots + \eta_M)$. In addition, if there are no external forces acting on the system, then $\sum_i \mathbf{F}_i = 0$ and three additional conservation laws exist of the form $\mathbf{K} = \mathbf{P} \exp(\eta_1)$ where $\mathbf{P} = \sum_i \mathbf{p}_i$ is the momentum of the centre of mass. In [38], it was shown how to prove, using the metric, the conservation laws, and the formalism outlined in section 2 that the NHC equations generate a proper canonical phase-space distribution in the physical coordinates and momenta. Note that if a separate NHC thermostat is coupled to each particle or even to each degree of freedom in the system, the \mathbf{K} conservation law no longer exists.

In order to develop a numerical integration algorithm for the NHC equations, it is important to keep in mind the modified Liouville theorem, equation (2.8). The complexity of the NHC equations is such that a Taylor series approach like that employed in [13] cannot be employed to derive a satisfactory integrator, i.e., one that does not lead to substantial drifts in the conserved energy [39]. Indeed, such an approach can be easily shown to violate equation (2.8). Thus, the NHC system is an example of a problem on which the power of the Liouville operator method can be brought to bear. We begin by writing the total Liouville operator for equations (4.1) as

$$iL = iL_1 + iL_2 + iL_T \quad (4.4)$$

where iL_1 and iL_2 are given by equation (3.8) and

$$iL_T = \sum_{k=1}^M \left[\frac{p_{\eta_k}}{Q_k} \frac{\partial}{\partial \eta_k} + G_k \frac{\partial}{\partial p_{\eta_k}} \right] - \sum_{i=1}^N \frac{p_{\eta_1}}{Q_1} \mathbf{p}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} - \sum_{k=1}^{M-1} \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \frac{\partial}{\partial p_{\eta_k}}. \quad (4.5)$$

The propagator is now factorized in a manner very similar to the velocity Verlet algorithm:

$$e^{iL\Delta t} = e^{iL_T\Delta t/2} e^{iL_2\Delta t/2} e^{iL_1\Delta t} e^{iL_2\Delta t/2} e^{iL_T\Delta t/2} + \mathcal{O}(\Delta t^3). \quad (4.6)$$

The only new feature in this scheme is the operator $\exp(iL_T\Delta t/2)$. Thus, the local error is still Δt^3 . Application of this operator to the phase space requires some care. Clearly, it needs to be further factorized into individual operators that can be applied analytically. However, the NHC equations constitute a set of differential equations with a time-scale separation between the particles and the thermostats and, therefore, a simple $\mathcal{O}(\Delta t^3)$ factorization scheme of the operator will not be accurate enough [19]. Thus, for this operator, a higher-order factorization is needed. Note that the overall integrator will still be $\mathcal{O}(\Delta t^3)$ despite the use of a higher-order method on the thermostat operator. The higher-order method we choose is the Suzuki–Yoshida (SY) scheme [44, 32], which involves the introduction of weighted time steps, $w_j\Delta t$, $j = 1, \dots, n_{\text{sy}}$, the value of n_{sy} determines the order of the method. The weights w_j are required to satisfy $\sum_{j=1}^{n_{\text{sy}}} w_j = 1$ and are chosen so as to cancel out the lower-order error

terms. Applying the SY scheme, the operator $\exp(iL_T\Delta t/2)$ becomes

$$e^{iL_T\Delta t/2} = \prod_{j=1}^{n_{sy}} [e^{iL_T w_j \Delta t/2}]. \quad (4.7)$$

The need to choose a large value of n_{sy} can be avoided by simply cutting the time step by a factor of n_c and applying the operator in equation (4.7) n_c times, i.e.

$$e^{iL_T\Delta t/2} = \prod_{i=1}^{n_c} \prod_{j=1}^{n_{sy}} [e^{iL_T w_j \Delta t/2n_c}]. \quad (4.8)$$

In this way, both n_c and n_{sy} can be adjusted so as to minimize the number of operations needed for satisfactory performance of the overall integrator. Having introduced the above scheme, it only remains to specify a particular factorization of the operator $\exp(iL_T w_j \Delta t/2n_c)$. Defining $\delta_j = w_j \Delta t/n_c$, we choose the following factorization [19]:

$$\begin{aligned} \exp\left[iL_T \frac{\delta_j}{2}\right] &= \exp\left[\frac{\delta_j}{4} G_M \frac{\partial}{\partial p_{\eta_M}}\right] \prod_{k=M-1}^1 \left[\exp\left[-\frac{\delta_j}{8} \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \frac{\partial}{\partial p_{\eta_k}}\right] \exp\left[\frac{\delta_j}{4} G_k \frac{\partial}{\partial p_{\eta_k}}\right] \right. \\ &\quad \times \exp\left[-\frac{\delta_j}{8} \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \frac{\partial}{\partial p_{\eta_k}}\right] \prod_{i=1}^N \exp\left[-\frac{\delta_j}{2} \frac{p_{\eta_i}}{Q_i} \mathbf{p}_i \cdot \frac{\partial}{\partial \mathbf{p}_i}\right] \\ &\quad \times \prod_{k=1}^M \exp\left[-\frac{\delta_j}{2} \frac{p_{\eta_k}}{Q_k} \frac{\partial}{\partial \eta_k}\right] \prod_{k=1}^{M-1} \left[\exp\left[-\frac{\delta_j}{8} \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \frac{\partial}{\partial p_{\eta_k}}\right] \right. \\ &\quad \left. \times \exp\left[\frac{\delta_j}{4} G_k \frac{\partial}{\partial p_{\eta_k}}\right] \exp\left[-\frac{\delta_j}{8} \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \frac{\partial}{\partial p_{\eta_k}}\right] \right] \exp\left[\frac{\delta_j}{4} G_M \frac{\partial}{\partial p_{\eta_M}}\right]. \quad (4.9) \end{aligned}$$

Although the overall scheme may seem complicated, the use of the direct translation technique simplifies the job of coding the algorithm. All of the operators appearing in equation (4.9) are either translation operators or operators of the form $\exp(cx\partial/\partial x)$, the action of which is

$$\exp\left[cx \frac{\partial}{\partial x}\right] x = x e^c. \quad (4.10)$$

We call such operators *scaling operators*, because the effect is to multiply x by an x -independent factor e^c . We note, finally, that multiple time-step integration can be incorporated into the NHC scheme by replacing the central three operators in equation (4.6) with the decomposition in equation (3.14) [19].

5. Integrating the MTK equations for the *NPT* ensemble

Having discussed the NHC equations of motion for generating the *NVT* ensemble, we are now in a position to introduce the MTK [17] equations for generating the *NPT* ensemble and a geometric scheme for integrating them. Although a geometric integrator for the MTK equations was introduced in [19], this scheme deviates slightly from the requirements of equation (2.8) and leads to other technical difficulties owing to the manner in which the classical propagator was factorized. Thus, here we present a scheme that exactly satisfies equation (2.8), reduces the technical complexities, and is considerably simpler to implement than that presented in [19].

The MTK equations allow the volume V to evolve dynamically while additionally providing a ‘barostat’ control to drive the fluctuations in the internal pressure estimator

$$P_{\text{int}} = \frac{1}{3V} \left[\sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} + \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i - 3V \frac{\partial U}{\partial V} \right] \quad (5.1)$$

in addition to a thermostat control that drives the kinetic energy fluctuations. For a system with no constraints, the equations of motion take the form

$$\begin{aligned} \dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} + \frac{p_\epsilon}{W} \mathbf{r}_i & \dot{\mathbf{p}}_i &= \mathbf{F}_i - \left(1 + \frac{1}{N}\right) \frac{p_\epsilon}{W} \mathbf{p}_i - \frac{p_{\eta_1}}{Q_1} \mathbf{p}_i \\ \dot{V} &= \frac{dV}{W} p_\epsilon & \dot{p}_\epsilon &= 3V(P_{\text{int}} - P) + \frac{1}{N} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - \frac{p_{\xi_1}}{Q'_1} p_\epsilon \\ \dot{\eta}_k &= \frac{p_{\eta_k}}{Q_k} \quad k = 1, \dots, M & \dot{p}_{\eta_k} &= G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \\ \dot{p}_{\eta_M} &= G_M & \dot{\xi}_k &= \frac{p_{\xi_k}}{Q'_k} \quad k = 1, \dots, M \\ \dot{p}_{\xi_k} &= G'_k - \frac{p_{\xi_{k+1}}}{Q'_{k+1}} p_{\xi_k} & \dot{p}_{\xi_M} &= G'_M. \end{aligned} \quad (5.2)$$

In equations (5.2), the variable p_ϵ with mass parameter W (having units of energy \times time²) corresponds to the barostat, coupling both to the positions and the momenta. Moreover, note that two Nosé–Hoover chains are coupled to the system, one to the particles and the other to the barostat. This device is particularly important, as the barostat tends to evolve on a much slower time scale than the particles. The heat-bath forces G'_k are defined by

$$G'_1 = \frac{p_\epsilon^2}{W} - kT \quad G'_k = \frac{p_{\xi_{k-1}}^2}{Q'_{k-1}} - kT. \quad (5.3)$$

The MTK equations have the conserved energy

$$H' = H(\mathbf{p}, \mathbf{r}) + \frac{p_\epsilon^2}{2W} + PV + \sum_{k=1}^M \left(\frac{p_{\eta_k}^2}{2Q_k} + \frac{p_{\xi_k}^2}{2Q'_k} \right) + 3NkT\eta_1 + kT \sum_{k=2}^M \eta_k + kT \sum_{k=1}^M \xi_k \quad (5.4)$$

and a phase-space metric factor

$$\sqrt{g(\mathbf{x})} = \exp \left(3N\eta_1 + \sum_{k=2}^M \eta_k + \sum_{k=1}^M \xi_k \right). \quad (5.5)$$

In addition, if $\sum_i \mathbf{F}_i = 0$, then three additional conservation laws of the form $\mathbf{K} = \exp[(1 + 1/N)\epsilon + \eta_1]$, where $\epsilon = (1/3) \ln(V/V_0)$. In order to prove that the MTK equations generate a correct isothermal–isobaric distribution, one needs to substitute equations (5.5) and the conservation laws into equation (2.10) and perform the integrals over all of the heat-bath variables and p_ϵ following the same procedure as was done for the canonical ensemble [38]. Again, if a separate NHC thermostat is coupled to each particle or to each degree of freedom, then the \mathbf{K} conservation law no longer exists.

Integrating the MTK equations is only slightly more difficult than integrating the NHC equations and builds on the technology already developed. Introducing, again, the variable $\epsilon = (1/3) \ln(V/V_0)$ and writing the total Liouville operator as

$$iL = iL_1 + iL_2 + iL_{\epsilon,1} + iL_{\epsilon,2} + iL_{\text{T-baro}} + iL_{\text{T-part}} \quad (5.6)$$

where

$$\begin{aligned} iL_1 &= \sum_{i=1}^N \left[\frac{\mathbf{p}_i}{m_i} + \frac{p_\epsilon}{W} \mathbf{r}_i \right] \cdot \frac{\partial}{\partial \mathbf{r}_i} & iL_2 &= \sum_{i=1}^N \left[\mathbf{F}_i - \alpha \frac{p_\epsilon}{W} \mathbf{p}_i \right] \cdot \frac{\partial}{\partial \mathbf{p}_i} \\ iL_{\epsilon,1} &= \frac{p_\epsilon}{W} \frac{\partial}{\partial \epsilon} & iL_{\epsilon,2} &= G_\epsilon \frac{\partial}{\partial p_\epsilon} \end{aligned} \quad (5.7)$$

and $iL_{\text{T-part}}$ and $iL_{\text{T-baro}}$ are defined in an analogous manner to equation (4.5). In equation (5.7), $\alpha = 1 + 1/N$, and

$$G_\epsilon = \alpha \sum_i \frac{\mathbf{p}_i^2}{m_i} + \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i - 3V \frac{\partial U}{\partial V} - PV. \quad (5.8)$$

The propagator is factorized in a manner that bears a very close resemblance to that of the NHC equations, namely

$$\begin{aligned} \exp(iL\Delta t) &= \exp\left(iL_{\text{T-baro}} \frac{\Delta t}{2}\right) \exp\left(iL_{\text{T-part}} \frac{\Delta t}{2}\right) \exp\left(iL_{\epsilon,2} \frac{\Delta t}{2}\right) \exp\left(iL_2 \frac{\Delta t}{2}\right) \\ &\quad \times \exp(iL_{\epsilon,1} \Delta t) \exp(iL_1 \Delta t) \exp\left(iL_2 \frac{\Delta t}{2}\right) \exp\left(iL_{\epsilon,2} \frac{\Delta t}{2}\right) \\ &\quad \times \exp\left(iL_{\text{T-part}} \frac{\Delta t}{2}\right) \exp\left(iL_{\text{T-baro}} \frac{\Delta t}{2}\right) + \mathcal{O}(\Delta t^3). \end{aligned} \quad (5.9)$$

In evaluating the action of this propagator, the Suzuki–Yoshida decomposition already developed for the NHC equations, together with equation (4.9) is applied to the operators $\exp(iL_{\text{T-baro}} \Delta t/2)$ and $\exp(iL_{\text{T-part}} \Delta t/2)$. The operators $\exp(iL_{\epsilon,1} \Delta t)$ and $\exp(iL_{\epsilon,2} \Delta t/2)$ are simple translation operators. The operators $\exp(iL_1 \Delta t)$ and $\exp(iL_2 \Delta t/2)$ are somewhat more complicated than their microcanonical or canonical ensemble counterparts due to the barostat coupling. The action of the operator $\exp(iL_1 \Delta t)$ can be determined by solving the differential equation

$$\dot{\mathbf{r}}_i = \mathbf{v}_i + v_\epsilon \mathbf{r}_i \quad (5.10)$$

for constant $\mathbf{v}_i = \mathbf{p}_i/m_i$ and constant $v_\epsilon = p_\epsilon/W$ for an arbitrary initial condition $\mathbf{r}_i(0)$ and evaluating the solution at $t = \Delta t$. This yields the evolution

$$\mathbf{r}_i(\Delta t) = \mathbf{r}_i(0) e^{v_\epsilon \Delta t} + \Delta t \mathbf{v}_i e^{v_\epsilon \Delta t/2} \frac{\sinh(v_\epsilon \Delta t/2)}{v_\epsilon \Delta t/2}. \quad (5.11)$$

Similarly, the action of $\exp(iL_2 \Delta t/2)$ can be determined by solving the differential equation

$$\dot{\mathbf{v}}_i = \frac{\mathbf{F}_i}{m_i} - \alpha v_\epsilon \mathbf{v}_i \quad (5.12)$$

for constant \mathbf{F}_i and an arbitrary initial condition $\mathbf{v}_i(0)$ and evaluating the solution at $t = \Delta t/2$. This yields the evolution

$$\mathbf{v}_i(\Delta t/2) = \mathbf{v}_i(0) e^{-\alpha v_\epsilon \Delta t/2} + \frac{\Delta t}{2m_i} \mathbf{F}_i e^{-\alpha v_\epsilon \Delta t/4} \frac{\sinh(\alpha v_\epsilon \Delta t/4)}{\alpha v_\epsilon \Delta t/4}. \quad (5.13)$$

In practice, the factor $\sinh(x)/x$ should be evaluated by a power series for small x to avoid numerical instabilities. These equations together with the Suzuki–Yoshida factorization of the thermostat operators completely define an integrator for the isothermal–isobaric ensemble that can be shown to satisfy equation (2.8). The integrator can be easily coded using the direct translation technique.

The new *NPT* integrator can also be applied within the r-RESPA framework. For two time steps, δt and $\Delta t = n\delta t$, the following contributions to the total Liouville operator are defined,

$$\begin{aligned} iL_1 &= \sum_{i=1}^N \left[\frac{\mathbf{p}_i}{m_i} + \frac{p_\epsilon}{W} \mathbf{r}_i \right] \cdot \frac{\partial}{\partial \mathbf{r}_i} & iL_2^{(1)} &= \sum_{i=1}^N \left[\mathbf{F}_i^{(1)} - \alpha \frac{p_\epsilon}{W} \mathbf{p}_i \right] \cdot \frac{\partial}{\partial \mathbf{p}_i} \\ iL_2^{(2)} &= \sum_{i=1}^N \mathbf{F}_i^{(2)} \cdot \frac{\partial}{\partial \mathbf{p}_i} & iL_{\epsilon,1} &= \frac{p_\epsilon}{W} \frac{\partial}{\partial \epsilon} \\ iL_{\epsilon,2}^{(1)} &= G_\epsilon^{(1)} \frac{\partial}{\partial p_\epsilon} & iL_{\epsilon,2}^{(2)} &= G_\epsilon^{(2)} \frac{\partial}{\partial p_\epsilon} \end{aligned} \quad (5.14)$$

where superscripts 1 and 2 are for the fast and slow contributions, respectively and,

$$G_\epsilon^{(1)} = \alpha \sum_i \frac{\mathbf{p}_i^2}{m_i} + \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i^{(1)} - 3V \frac{\partial U^{(1)}}{\partial V} - 3P^{(1)}V \quad (5.15)$$

$$G_\epsilon^{(2)} = \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i^{(2)} - 3V \frac{\partial U^{(2)}}{\partial V} - 3P^{(2)}V. \quad (5.16)$$

The variables $P^{(1)}$ and $P^{(2)}$ should be chosen such that the external pressure $P = P^{(1)} + P^{(2)}$. (In the next section, three different possible ways of subdividing the applied external pressure will be investigated.) The factorization scheme for the propagator is

$$\begin{aligned} \exp(iL\Delta t) &= \exp\left(iL_{\text{T-baro}} \frac{\Delta t}{2}\right) \exp\left(iL_{\text{T-part}} \frac{\Delta t}{2}\right) \exp\left(iL_{\epsilon,2}^{(2)} \frac{\Delta t}{2}\right) \exp\left(iL_2^{(2)} \frac{\Delta t}{2}\right) \\ &\times \left[\exp\left(iL_{\epsilon,2}^{(1)} \frac{\delta t}{2}\right) \exp\left(iL_2^{(1)} \frac{\delta t}{2}\right) \exp(iL_{\epsilon,1}\delta t) \exp(iL_1\delta t) \right. \\ &\times \left. \exp\left(iL_2^{(1)} \frac{\delta t}{2}\right) \exp\left(iL_{\epsilon,2}^{(1)} \frac{\delta t}{2}\right) \right]^n \exp\left(iL_2^{(2)} \frac{\Delta t}{2}\right) \exp\left(iL_{\epsilon,2}^{(2)} \frac{\Delta t}{2}\right) \\ &\times \exp\left(iL_{\text{T-part}} \frac{\Delta t}{2}\right) \exp\left(iL_{\text{T-baro}} \frac{\Delta t}{2}\right) + \mathcal{O}(\Delta t^3). \end{aligned} \quad (5.17)$$

Note that, because G_ϵ depends on the forces, \mathbf{F}_i , it is necessary to update both the particles and the barostat in the reference system.

6. Illustration and analysis of the new *NPT* integrator

In order to illustrate the new *NPT* integrator and analyse its structure, we consider the simple example of a single particle with coordinate q and momentum p moving in a one-dimensional periodic potential of the form

$$U(q) = \frac{m\omega^2 V^2}{4\pi^2} \left[1 - \cos\left(\frac{2\pi q}{V}\right) \right]. \quad (6.1)$$

Here, V is the volume of the one-dimensional box. The *NPT* equations of motion for this problem are integrated using a time step of $\Delta t = 0.05$. Other parameters are chosen to be $n_c = 4$, $n_{\text{sy}} = 3$, $M = 2$, $m = 1$, $\omega = 1$, $W = 18$, $kT = 1$, $P = 1$, $Q_k = 1$, and $Q'_k = \sqrt{\frac{1}{20}}$. Figure 1 shows the position and volume distributions generated from the simulation together with their analytic values. It can be seen from the figure that these distributions are perfectly generated.

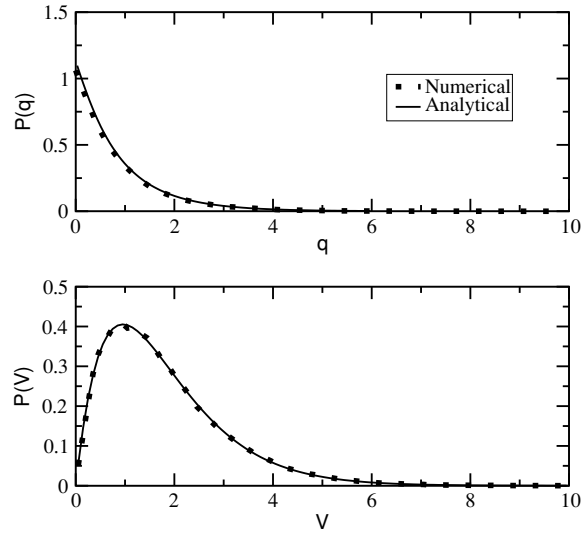


Figure 1. Comparison of the position (top) and volume (bottom) distribution functions from the numerical simulation of equations (5.2) integrating using equation (5.9).

In the appendix, we perform an analysis of the Jacobian of the new *NPT* integrator, calculating it both analytically and using the above example to demonstrate numerically that the proper phase-space Jacobian is generated and, therefore, that the generalized Liouville theorem in equation (2.8) is obeyed.

7. Additional examples

The new *NPT* algorithm is also applied to study two realistic models using molecular dynamics simulations: a simple Lennard–Jones (LJ) fluid and a fluid of flexible diatomic molecules interacting via the LJ potential. In the latter case, the multiple time scale r-RESPA method [40] is employed.

7.1. Lennard–Jones fluid

The simulations were performed using reduced units ($\epsilon/k = 119.8$ K, $\sigma = 3.405$ and $m = 39.95$ g mol⁻¹). The system contains $N = 256$ particles in a cell with periodic boundary conditions (PBC) in all directions at $T^* = 1.4$ and $P^* = 1.279$. This state is above the critical point. The pressure was obtained from an *NVT* simulation at $\rho^* = 0.7$ using the same number of particles [16]. The dispersion interactions were calculated using the lattice sum method or Ewald-type summation applied to $1/r^6$ interactions. Thus, the results are independent of the cut-off distance. The simulations were performed using a reduced time step $\Delta t^* = 0.002$. In addition to a thermostat on p_ϵ , a separate NHC thermostat was applied to each particle in the fluid. Thus, the thermostat and barostat time scale parameters W , Q_p and Q_b were determined by the recommended formula [18, 19]:

$$\begin{aligned}
 Q_p^1 &= Q_1 = 3kT\tau_p^2 & Q_p^k &= Q_k = kT\tau_p^2 \\
 Q_b^1 &= Q'_1 = kT\tau_b^2 & Q_b^k &= Q'_k = kT\tau_b^2 \\
 W &= (3N + 3)kT\tau_b^2
 \end{aligned} \tag{7.1}$$

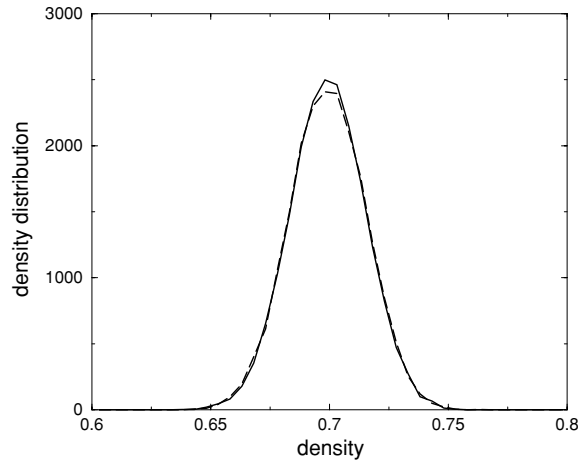


Figure 2. Density distribution for LJ fluids for two set of (τ_p^*, τ_b^*) . The solid line is for (0.2,0.2) and dashed line for (2.0,0.2).

where τ_p and τ_b are characteristic time scales of the particles and barostat, respectively. Typically, τ_p and τ_b can be chosen directly proportional to the time step Δt . In the present simulations, two sets of thermostat and barostat parameters were used to evaluate the error in the conserved quantity equation (5.4), denoted $\Delta E = \frac{1}{N_c} \sum_i |(E_i - E_0)/E_0|$, where i runs over the number of configurations N_c . The first set of parameters was $\tau_p^* = \tau_b^* = 100\Delta t^*$ for which it was found that $\Delta E = 4.2 \times 10^{-5}$ over a run of length two million steps. By changing τ_b^* or τ_p^* to $1000\Delta t^*$ an improvement in the energy conservation was observed with $\Delta E = 5.0 \times 10^{-6}$ and $\Delta E = 4.7 \times 10^{-7}$, respectively. No drift in the conserved quantity was observed in any of the simulations.

The systems were equilibrated for 5×10^5 time steps and additional 2×10^6 configurations were used to obtain the average properties. The average kinetic energy and pressure fluctuations were found to converge in a satisfactory manner to the imposed temperature T and pressure P . In all simulations, the average density was $\langle \rho^* \rangle = 0.70 \pm 0.02$ and the average potential energy $\langle U^* \rangle = -4.63 \pm 0.12$ (compared with $\langle U^* \rangle = -4.634 \pm 0.009$ from NVT simulations [16]). These results show that both ensembles yield equivalent equilibrium thermodynamics. Figure 2 shows the density distributions obtained in the NPT simulations using two set of parameters (τ_p^*, τ_b^*) , (0.2, 0.2) and (2.0, 0.2). Although the errors were different, the density distributions are seen to be nearly the same. The slight difference between the distributions most likely indicate the need for somewhat longer runs.

7.2. A LJ fluid of flexible diatomic molecules

The next illustrative example is a fluid of flexible diatomic molecules interacting via a LJ potential. Each diatomic molecule contains two equal-sized atoms with a harmonic bond potential $U_b(r) = k_r(r - r_0)^2/2$, where r is the distance between particles, k_r is the spring constant and r_0 is the equilibrium distance. The system contains $N = 256$ molecules in a box with PBCs at supercritical conditions. The atoms of different molecules interact via a LJ potential with parameters $\epsilon/k = 104$ K and $\sigma = 3.91$ Å; these values are similar to those used to simulate hydrocarbons. The simulations were performed in reduced units. The cut-off distance, $r_c^* = r_c/\sigma$ of 4 was used and the force $\mathbf{F}(r)$ was switched smoothly from $r_c^* - r_s^* = 3.7$ to r_c^* as recommended in [37] by multiplying the force by a switching function $S(r)$. The use

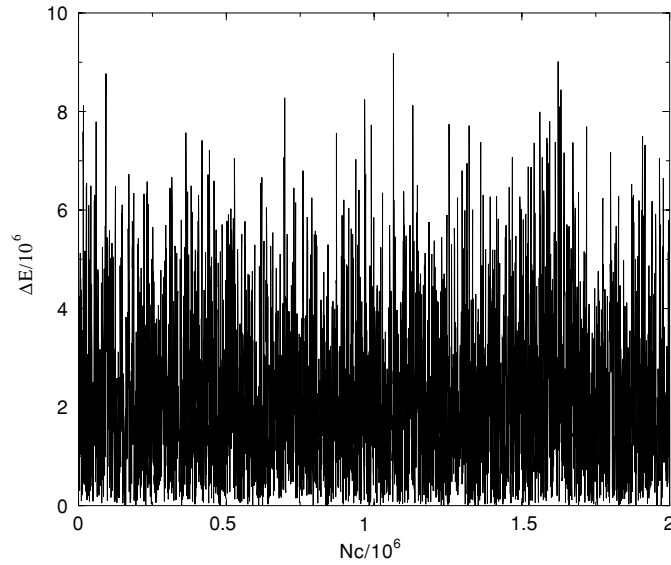


Figure 3. Error in the conserved quantity as a function of number of configurations for the fluid of flexible diatomics using $n = 4$, $P^{(1)} = nP/(n + 1)$ and $P^{(2)} = P/(n + 1)$.

of a switch eliminates possible discontinuities in the pressure estimator due to the finite-size cut-off radius and allows a constant cut-off to be used rather than a volume-dependent cut-off. The switching function is given by $S(r) = 1 + \lambda^2(2\lambda - 3)$, where $\lambda = (r - (r_c - r_s)/r_s)$. The other parameters were taken to be $k_r^* = 846.5$, $r_0^* = 0.394$, $\Delta t^* = 0.0001$, $\tau_p^* = 20\Delta t^*$ and $\tau_b^* = 100\Delta t^*$. The external temperature and pressure were fixed to $T^* = 5.0$ and $P^* = 1.5$. The system was equilibrated for 2×10^5 time steps and the average properties were calculated for additional 2×10^6 steps.

The first simulation was carried out using a single time step and yielded an average density of $\rho^* = 0.569$ and configurational energy of $U^* = -1.851$. During this simulation, the average of $P^{(1)} = (\alpha \sum_i m_i \mathbf{v}_i^2 + \sum_i \mathbf{F}_i^{(1)} \cdot \mathbf{r}_i)/3V$ and $P^{(2)} = \sum_i \mathbf{F}_i^{(2)} \cdot \mathbf{r}_i/3V$ were calculated, where $\mathbf{F}^{(1)}$ represents the intramolecular harmonic force and $\mathbf{F}^{(2)}$ represents the intermolecular LJ force. The calculated values were $P^{(1)*} = 1.79$ and $P^{(2)*} = -0.29$, respectively. As expected, the sum of these two estimators $P^* = P^{(1)*} + P^{(2)*}$ yields the applied external pressure of 1.5.

The fluid of flexible diatomics is a good test system for the multiple time-scale NPT method. Here, a small time step δt^* is chosen for the fast intramolecular forces $\mathbf{F}^{(1)}$ and a larger time step Δt^* is chosen the slower intermolecular LJ interactions $\mathbf{F}^{(2)}$. Here, δt^* can be chosen to be the time step 0.0001 used in the above single time-step simulation, while $\Delta t^* = n\delta t^*$ is chosen to be a factor n times larger. The thermostat and barostat parameters had the same ratio with Δt^* as in the simulation with one time step. To perform the simulations using different time steps, it is necessary to calculate $G_\epsilon = \alpha \sum_i m_i \mathbf{v}_i^2 + g^{(1)} + ng^{(2)}$ where $g^{(1)} = \sum_i \mathbf{F}_i^{(1)} \cdot \mathbf{r}_i - 3P^{(1)}V$ and $g^{(2)} = \sum_i \mathbf{F}_i^{(2)} \cdot \mathbf{r}_i - 3P^{(2)}V$. The term $ng^{(2)}$ is only added when the slow forces are evaluated. Three options were analysed to choose $P^{(1)*}$ and $P^{(2)*}$: (A) $P^{(1)*} = P^*$ and $P^{(2)*} = 0$, (B) $P^{(1)*} = nP^*/(n + 1)$ and $P^{(2)*} = P^*/(n + 1)$ and (C) $P^{(1)*} = 1.79$ and $P^{(2)*} = -0.29$. Table 1 gives the simulation results for these three options using $n = 4$. Figure 3 shows the error in the conserved quantity for option B. The three simulations conserve the energy in equation (5.4) without a drift and give essentially the same results as those obtained from a single time step. The option B might be better in the simulation

Table 1. Results for a fluid of flexible diatomic molecules obtained using the new *NPT* integrator. Here, the calculated temperature, pressure, potential energy, density and energy conservation measure ΔE are reported.

System	A	B	C
T^*	5.0	5.0	5.0
$\langle P_{\text{int}}^* \rangle$	1.5	1.5	1.5
$\langle U^* \rangle$	-1.853	-1.853	-1.851
$\langle \rho^* \rangle$	0.571	0.570	0.571
ΔE	3.1×10^{-6}	2.9×10^{-6}	2.7×10^{-6}

Table 2. Results for a fluid of flexible diatomic molecules obtained using the new *NPT* integrator with two time steps for different values of n . $P^{(1)*} = nP^*/(n+1)$ and $P^{(2)*} = P^*/(n+1)$.

n	1	2	4	8
T^*	5.0	5.0	5.0	5.0
$\langle P_{\text{int}}^* \rangle$	1.5	1.5	1.5	1.5
$\langle U^* \rangle$	-1.851	-1.843	-1.853	-1.848
$\langle \rho^* \rangle$	0.569	0.569	0.570	0.570
ΔE	2.8×10^{-6}	2.9×10^{-6}	2.9×10^{-6}	3.2×10^{-6}

of molecular fluids using r-RESPA because the external pressure is split between the different force contributions. Table 2 gives results for different values of n using the option B.

8. Conclusions

The isothermal–isobaric (*NPT*) molecular dynamics approach has been re-examined for the purpose of developing a new measure-preserving, reversible geometric integrator that improves on previously introduced integrators. In the context of developing the new approach, we have reviewed the underlying statistical mechanical theory of non-Hamiltonian phase spaces, the Liouville operator formalism and the symplectic property of Hamiltonian systems. It was pointed out that gaps still exist in the generalization of Hamiltonian concepts to non-Hamiltonian systems.

The new *NPT* integrator was shown to satisfy the volume conservation law of equation (2.8) and was observed to perform well on realistic problems. The method was also extended to include multiple time steps for systems with motion on several time scales. Future work will include generalizing the scheme for fully flexible cell fluctuations, incorporation of holonomic constraints into the scheme, substitution of thermal control methods other than the Nosé–Hoover chain approach used here, and incorporation of the new approach into the multiple time-step isokinetic scheme recently introduced by Minary *et al* [20].

Acknowledgments

JA thanks Mauricio Romero for helpful discussions and to CONACyT for financial support. RLR thanks CONACyT for a scholarship. MET acknowledges NSF CHE-0121375 and NSF CHE-0310107.

Appendix. Analysis of the Jacobian

In this appendix the analytical derivation of the Jacobian of transformation for a potential $U(q)$ with the force $F(q) = -dU/dq$ is presented. The analysis is based on the factorization

scheme of equation (5.9) with equation (4.9) used to factorize the thermostat operators. For simplicity, we set $n_{\text{sy}} = n_c = 1$ in the Suzuki–Yoshida scheme; however, it does not affect the basic conclusion. We also choose the length M of Nosé–Hoover chains to be 2 for concreteness. Finally, for further simplicity, we set $Q_1 = Q_2 = Q'_1 = Q'_2 = 1$.

With these parameters, the total number of operators in the factorization is 50; however, if the scaling and translation operations in equation (4.9) are combined into one, then the total number of operators reduces to 34. Finally, the phase-space vector \mathbf{x} contains the 12 variables

$$\mathbf{x} = (p, p_{\eta_1}, p_{\eta_2}, p_{\xi_1}, p_{\xi_2}, p_\epsilon, q, \eta_1, \eta_2, \xi_1, \xi_2, V). \quad (\text{A.1})$$

Each operator can be regarded as a specific transformation on the phase space, and thus, the product of operators produces a series of transformations

$$\mathbf{x}(0) \rightarrow \mathbf{x}(1) \rightarrow \cdots \mathbf{x}(34) \quad (\text{A.2})$$

where $\mathbf{x}(34) = \mathbf{x}(\Delta t)$. The overall Jacobian matrix of the full transformation produced by the factorization of the one-time-step propagator is the product of the Jacobian matrices of each individual transformation step

$$J_{\text{tot}} = \prod_{\alpha=1}^{34} J_\alpha \quad (\text{A.3})$$

where,

$$J_\alpha^{ij} = \frac{\partial x^i(\alpha)}{\partial x^j(\alpha-1)}. \quad (\text{A.4})$$

Consequently, the total Jacobian determinant will be a product of the determinants of the individual Jacobian matrices

$$\det(J_{\text{tot}}) = \prod_{\alpha=1}^{34} \det(J_\alpha). \quad (\text{A.5})$$

Since each individual step changes only one variable at a time, the individual Jacobian matrices are simple to calculate. In fact, they will differ from identity matrices by only one element for each transformation. For example, the first variable transformed is p_{ξ_2} , which transforms according to

$$p_{\xi_2}(1) = p_{\xi_2}(0) + \frac{\Delta t}{4} G_{\xi_2}(0). \quad (\text{A.6})$$

All other variables remain unchanged. Since G_{ξ_2} depends on p_{ξ_1} , the only element of J_1 that will differ from the identity matrix is the J_1^{12} element. This element is given by

$$J_1^{12} = \frac{\Delta t}{2} p_{\xi_1}(0) \quad (\text{A.7})$$

and $\det(J_1) = 1$.

The second variable transformed in the steps $\mathbf{x}(1) \rightarrow \mathbf{x}(2)$ is p_{ξ_1} , which transforms as

$$p_{\xi_1}(2) = p_{\xi_1}(1) e^{-\Delta t p_{\xi_2}(1)/4} + \frac{\Delta t}{4} G_{\xi_1} e^{-\Delta t p_{\xi_2}(1)/8} \quad (\text{A.8})$$

(recall that scaling and translation operators have been combined). All other variables retain their values from the previous step. Thus, the Jacobian matrix of this transformation has only two elements that differ from the identity matrix, J_2^{21} and J_2^{22} . These elements are given by

$$\begin{aligned} J_2^{21} &= -\frac{\Delta t}{4} p_{\xi_1}(1) e^{-\Delta t p_{\xi_2}(1)/4} - \frac{\Delta t^2}{32} G_{\xi_1}(1) e^{-\Delta t p_{\xi_2}(1)/8} \\ J_2^{22} &= e^{-\Delta t p_{\xi_2}(1)/4}. \end{aligned} \quad (\text{A.9})$$

Hence, the Jacobian of this matrix is easily seen to be

$$\det(J_2) = e^{-\Delta t p_{\xi_2}(1)/4}. \quad (\text{A.10})$$

Repeating this procedure 32 more times for each step of the integrator yields the overall Jacobian of the integrator:

$$\det(J_{\text{tot}}) = e^{-\Delta t(p_{\xi_2}(1)+p_{\xi_2}(5)+p_{\xi_2}(28)+p_{\xi_2}(32)+p_{\eta_2}(8)+p_{\eta_2}(12)+p_{\eta_2}(21)+p_{\eta_2}(25))/4} \\ \times e^{-\Delta t(p_{\xi_1}(4)+p_{\xi_1}(31)+p_{\eta_1}(11)+p_{\eta_1}(24))/2} e^{-\Delta t(\frac{p_{\epsilon}(15)}{W}-\epsilon(16)-\frac{p_{\epsilon}(17)}{W}+\frac{p_{\epsilon}(18)}{W})}. \quad (\text{A.11})$$

However, recognizing that

$$\begin{aligned} p_{\xi_2}(1) &= p_{\xi_2}(\Delta t/4) & p_{\xi_2}(5) &= p_{\xi_2}(\Delta t/4) \\ p_{\xi_2}(28) &= p_{\xi_2}(3\Delta t/4) & p_{\xi_2}(32) &= p_{\xi_2}(3\Delta t/4) \\ p_{\eta_2}(8) &= p_{\eta_2}(\Delta t/4) & p_{\eta_2}(12) &= p_{\eta_2}(\Delta t/4) \\ p_{\eta_2}(21) &= p_{\eta_2}(3\Delta t/4) & p_{\eta_2}(25) &= p_{\eta_2}(3\Delta t/4) \\ p_{\xi_1}(4) &= p_{\xi_1}(\Delta t/4) & p_{\xi_1}(31) &= p_{\xi_1}(3\Delta t/4) \\ p_{\eta_1}(11) &= p_{\eta_1}(\Delta t/4) & p_{\eta_1}(24) &= p_{\eta_1}(3\Delta t/4) \end{aligned} \quad (\text{A.12})$$

and,

$$\begin{aligned} \frac{1}{2}(\xi_2(\Delta t/2) - \xi_2(0)) &= \frac{\Delta t}{4} p_{\xi_2}(\Delta t/4) \\ \frac{1}{2}(\xi_2(\Delta t) - \xi_2(\Delta t/2)) &= \frac{\Delta t}{4} p_{\xi_2}(3\Delta t/4) \\ \xi_2(\Delta t) - \xi_2(\Delta t/2) + \xi_2(\Delta t/2) - \xi_2(0) &= \frac{\Delta t}{2}(p_{\xi_2}(\Delta t/4) + p_{\xi_2}(3\Delta t/4)) \\ \xi_2(\Delta t) - \xi_2(0) &= \frac{\Delta t}{2}(p_{\xi_2}(\Delta t/4) + p_{\xi_2}(3\Delta t/4)) \\ \frac{1}{2}(\eta_2(\Delta t/2) - \eta_2(0)) &= \frac{\Delta t}{4} p_{\eta_2}(\Delta t/4) \\ \frac{1}{2}(\eta_2(\Delta t) - \eta_2(\Delta t/2)) &= \frac{\Delta t}{4} p_{\eta_2}(3\Delta t/4) \\ \eta_2(\Delta t) - \eta_2(\Delta t/2) + \eta_2(\Delta t/2) - \eta_2(0) &= \frac{\Delta t}{2}(p_{\eta_2}(\Delta t/4) + p_{\eta_2}(3\Delta t/4)) \\ \eta_2(\Delta t) - \eta_2(0) &= \frac{\Delta t}{2}(p_{\eta_2}(\Delta t/4) + p_{\eta_2}(3\Delta t/4)) \\ \xi_1(\Delta t/2) - \xi_1(0) &= \frac{\Delta t}{2} p_{\xi_1}(\Delta t/4) \\ \xi_1(\Delta t) - \xi_1(\Delta t/2) &= \frac{\Delta t}{2} p_{\xi_1}(3\Delta t/4) \\ \xi_1(\Delta t) - \xi_1(\Delta t/2) + \xi_1(\Delta t/2) - \xi_1(0) &= \frac{\Delta t}{2}(p_{\xi_1}(\Delta t/4) + p_{\xi_1}(3\Delta t/4)) \\ \xi_1(\Delta t) - \xi_1(0) &= \frac{\Delta t}{2}(p_{\xi_1}(\Delta t/4) + p_{\xi_1}(3\Delta t/4)) \\ \eta_1(\Delta t/2) - \eta_1(0) &= \frac{\Delta t}{2} p_{\eta_1}(\Delta t/4) \\ \eta_1(\Delta t) - \eta_1(\Delta t/2) &= \frac{\Delta t}{2} p_{\eta_1}(3\Delta t/4) \\ \eta_1(\Delta t) - \eta_1(\Delta t/2) + \eta_1(\Delta t/2) - \eta_1(0) &= \frac{\Delta t}{2}(p_{\eta_1}(\Delta t/4) + p_{\eta_1}(3\Delta t/4)) \\ \eta_1(\Delta t) - \eta_1(0) &= \frac{\Delta t}{2}(p_{\eta_1}(\Delta t/4) + p_{\eta_1}(3\Delta t/4)). \end{aligned} \quad (\text{A.13})$$

Table 3. Numerical and analytical Jacobian.

Δt	$\det(J_N)$	$\det(J_A)$	$ \det(J_N) - \det(J_A) $
0.01	0.998 1491	0.998 1435	5.6×10^{-6}
0.03	0.983 0623	0.983 0591	3.2×10^{-6}
0.05	0.952 7543	0.952 7493	5.0×10^{-6}
0.07	0.907 9818	0.907 9821	3.0×10^{-7}
0.09	0.850 5255	0.850 5250	5.0×10^{-7}
0.10	0.817 8399	0.817 8414	1.5×10^{-6}

The Jacobian becomes

$$\det(J_{\text{tot}}) = e^{-(\xi_2(\Delta t) + \xi_1(\Delta t) + \eta_2(\Delta t) + \eta_1(\Delta t))} e^{(\xi_2(0) + \xi_1(0) + \eta_2(0) + \eta_1(0))} \times e^{-\Delta t(p_\epsilon(15)/W - \epsilon(16) - p_\epsilon(17)/W + p_\epsilon(18)/W)} \quad (\text{A.14})$$

where,

$$\frac{p_\epsilon(15)}{W} = \frac{p_\epsilon(17)}{W} = \frac{p_\epsilon(18)}{W} = \frac{p_\epsilon(\Delta t/2)}{W} \quad (\text{A.15})$$

and,

$$\epsilon(\Delta t) = \epsilon(0) + \frac{p_\epsilon(\Delta t/2)}{W}. \quad (\text{A.16})$$

However, recalling that

$$\epsilon = \ln \frac{V}{V(0)} \quad (\text{A.17})$$

where $V(0)$ is the initial volume, it follows that $\epsilon(0) = 0$. Thus,

$$\epsilon(\Delta t) = \frac{p_\epsilon(\Delta t/2)}{W}. \quad (\text{A.18})$$

Finally, the Jacobian of the transformation $x_t \rightarrow x_0$ is

$$\det(J_{\text{tot}}) = e^{-(\xi_2(\Delta t) + \xi_1(\Delta t) + \eta_2(\Delta t) + \eta_1(\Delta t))} e^{(\xi_2(0) + \xi_1(0) + \eta_2(0) + \eta_1(0))} \quad (\text{A.19})$$

which is precisely the Jacobian computed directly from the equations of motion. Thus, the new integrator satisfies the generalized Liouville theorem in equation (2.8).

In addition to the analytical derivation, a numerical calculation of the Jacobian was performed as well using the potential in equation (6.1). The numerical calculation was performed as follows: In order to calculate each element of the 12×12 Jacobian matrix, the initial value of each phase-space variable was displaced by amounts h and $-h$, with $h = 10^{-6}$, and the integrator was run for 1 step to produce the resultant phase-space vector at time Δt . A numerical derivative was then calculated in order to obtain each element in the matrix. Finally, the determinant of the matrix was then determined. Table 3 compares the values of numerical and analytical Jacobian for different-sized time steps.

It can be seen that the numerical and analytical Jacobian values agree to within the error of the numerical differentiation with no degradation as the time step is increased. This serves as an additional validation of the new integrator.

References

- [1] Andersen H C 1980 Molecular dynamics at constant temperature and/or pressure *J. Chem. Phys.* **72** 2384
- [2] Andrey L 1986 *Phys. Lett. A* **114** 183

- [3] Bond S D, Leimkuhler B J and Laird B B 1999 The Nosé–Poincaré method for constant temperature molecular dynamics *J. Comput. Phys.* **151** 114
- [4] Calvo M P and Sanz-Serna J M 1994 *Numerical Hamiltonian Problems* (London, UK: Chapman and Hall)
- [5] Creutz M and Goksch A 1989 *Phys. Rev. Lett.* **63** 9
- [6] Dubrovin B A, Fomenko A T and Novikov S P 1985 *Modern Geometry—Methods and Applications Part I* (New York: Springer)
- [7] Dubrovin B A, Fomenko A T and Novikov S P 1985 *Modern Geometry—Methods and Applications Part II* (New York: Springer)
- [8] Engle R D, Skeel R D and Drees M 2005 Monitoring energy drift with shadow Hamiltonians *J. Comput. Phys.* **206** 432
- [9] Hoover W G 1985 Canonical dynamics—equilibrium phase-space distributions *Phys. Rev. A* **31** 1695
- [10] Hoover W G 1986 Constant-pressure equations of motion *Phys. Rev. A* **34** 2499
- [11] Miller T F III, Eleftheriou M, Pattnaik P, Ndirango A, News D and Martyna G J 2003 Symplectic quaternion scheme for biophysical molecular dynamics *J. Chem. Phys.* **116** 8649
- [12] Izaguirre J A and Hampton S S 2004 Shadow hybrid Monte Carlo: an efficient propagator in the phase space of macromolecules *J. Comput. Phys.* **200** 481
- [13] Jang S and Voth G A 1997 Simple reversible molecular dynamics algorithms for Nose–Hoover chain dynamics *J. Chem. Phys.* **107** 9514
- [14] Leimkuhler B J and Sweet C R 2004 The canonical ensemble via symplectic integrators using Nose and Nose–Poincaré chains *J. Chem. Phys.* **121** 108
- [15] Liu Y and Tuckerman M E 2000 Generalized Gaussian moment thermostating: a new continuous dynamical approach to the canonical ensemble *J. Chem. Phys.* **112** 1685
- [16] Lopez-Lemus J and Alejandre J 2002 Thermodynamic and transport properties of simple fluids using lattice sums: bulk phases and liquid-vapour interface *Mol. Phys.* **100** 2983
- [17] Martyna G J, Tobias D J and Klein M L 1994 Constant-pressure molecular-dynamics algorithms *J. Chem. Phys.* **101** 4177
- [18] 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
- [19] Martyna G J, Tuckerman M E, Tobias D J and Klein M L 1996 Explicit reversible integrators for extended systems dynamics *Mol. Phys.* **87** 1117
- [20] Minary P, Tuckerman M E and Martyna G J 2004 Long time-step molecular dynamics for enhanced conformational sampling in biomolecular systems *Phys. Rev. Lett.* **93** 150201
- [21] Nosé S 1984 A unified formulation of the constant temperature molecular dynamics methods *J. Chem. Phys.* **81** 511
- [22] Ramshaw J D 1986 *Phys. Lett. A* **116** 110
- [23] Sardanashvily G 2002 The Lyapunov stability of first-order dynamic equations with respect to time-dependent Riemannian metrics *Preprint* nlin.CD/02010660v1
- [24] Sardanashvily G 2002 The Lyapunov stability of first-order dynamic equations with respect to time-dependent Riemannian metrics: an example *Preprint* nlin.CD/0203031v1
- [25] Schutz B 1987 *Geometrical Methods of Mathematical Physics* (Cambridge: Cambridge University Press)
- [26] Sergi A 2003 *Phys. Rev. E* **67** 021101
- [27] Sergi A 2004 *Phys. Rev. E* **69** 021109
- [28] Sergi A 2004 *Phys. Rev. E* **72** 031104
- [29] Sergi A and Ferrario M 2001 *Phys. Rev. E* **64** 056125
- [30] Skeel R D and Hardy D J 2001 Practical construction of modified Hamiltonians *SIAM J. Sci. Comput.* **23** 1172
- [31] Sturgeon J B and Laird B B 2000 Symplectic algorithm for constant-pressure molecular dynamics using a Nose–Poincaré thermostat *J. Chem. Phys.* **112** 3474
- [32] Suzuki M 1991 General-theory of fractal path-integrals with applications to many-body theories and statistical physics *J. Math. Phys.* **32** 400
- [33] Tarasov V E 2005 Phase-space metric for non-Hamiltonian systems *J. Phys. A: Math. Gen.* **38** 2145
- [34] Tarasov V E 2005 Stationary solutions of Liouville equations for non-Hamiltonian systems *Ann. Phys.* **316** 393
- [35] Trotter H F 1959 *Proc. Am. Math. Soc.* **10** 545
- [36] Tuckerman M E in preparation
- [37] Tuckerman M E, Hughes A and Martyna G J 1999 Molecular dynamics algorithm for path integrals at constant pressure *J. Chem. Phys.* **110** 3275
- [38] Tuckerman M E, Liu Y, Ciccotti G and Martyna G J 2001 Non-Hamiltonian molecular dynamics: generalizing Hamiltonian phase space principles to non-Hamiltonian systems *J. Chem. Phys.* **115** 1678

- [39] Tuckerman M E and Martyna G J 1999 Comment on simple reversible molecular dynamics algorithms for Nose–Hoover chain dynamics *J. Chem. Phys.* **110** 3623
- [40] Tuckerman M E, Martyna G J and Berne B J 1992 Reversible multiple time scale molecular dynamics *J. Chem. Phys.* **97** 1990
- [41] Tuckerman M E, Mundy C J and Klein M L 1997 Toward a statistical thermodynamics of steady states *Phys. Rev. Lett.* **78** 2042
- [42] Tuckerman M E, Mundy C J and Martyna G J 1999 On the classical statistical mechanics of non-Hamiltonian systems *Europhys. Lett.* **45** 149
- [43] Wilcox R M 1967 *J. Math. Phys.* **8** 962
- [44] Yoshida H 1990 Construction of higher-order symplectic integrators *Phys. Lett. A* **150** 262