

Review

# Holonomic Constraints: A Case for Statistical Mechanics of Non-Hamiltonian Systems

Giovanni Ciccotti <sup>1,2,3</sup>  and Mauro Ferrario <sup>4,\*</sup> 

<sup>1</sup> Institute for Applied Mathematics “Mauro Picone” (IAC), CNR, Via dei Taurini 19, 00185 Rome, Italy; giovanni.ciccotti@roma1.infn.it

<sup>2</sup> University of Rome “La Sapienza”, P.le Aldo Moro 5, 00185 Rome, Italy

<sup>3</sup> University College Dublin (UCD), Belfield, Dublin 4, Ireland

<sup>4</sup> Dipartimento di Scienze Fisiche, Informatiche e Matematiche, University of Modena and Reggio Emilia, Via Campi 213/A, 41125 Modena, Italy

\* Correspondence: mauro.ferrario@unimore.it; Tel.: +39-05-9205-5291

Received: 8 January 2018; Accepted: 25 January 2018; Published: 1 February 2018

**Abstract:** A dynamical system submitted to holonomic constraints is Hamiltonian only if considered in the reduced phase space of its generalized coordinates and momenta, which need to be defined *ad hoc* in each particular case. However, specially in molecular simulations, where the number of degrees of freedom is exceedingly high, the representation in generalized coordinates is completely unsuitable, although conceptually unavoidable, to provide a rigorous description of its evolution and statistical properties. In this paper, we first review the state of the art of the numerical approach that defines the way to conserve exactly the constraint conditions (by an algorithm universally known as SHAKE) and permits integrating the equations of motion directly in the phase space of the natural Cartesian coordinates and momenta of the system. We then discuss in detail SHAKE numerical implementations in the notable cases of Verlet and velocity-Verlet algorithms. After discussing in the same framework how constraints modify the properties of the equilibrium ensemble, we show how, at the price of moving to a dynamical system no more (directly) Hamiltonian, it is possible to provide a direct interpretation of the dynamical system and so derive its Statistical Mechanics both at equilibrium and in non-equilibrium conditions. To achieve that, we generalize the statistical treatment to systems no longer conserving the phase space volume (equivalently, we introduce a non-Euclidean invariant measure in phase space) and derive a generalized Liouville equation describing the ensemble even out of equilibrium. As a result, we can extend the response theory of Kubo (linear and nonlinear) to systems subjected to constraints.

**Keywords:** holonomic constraints; non-Hamiltonian dynamics; SHAKE

## 1. Introduction

The dynamical and statistical behavior of a mechanical system of many degrees of freedom subjected to holonomic constraints presents specific features that seem worth presenting and discussing in a unified framework. A mechanical Hamiltonian system is a system whose evolution is derivable from a standard Hamiltonian

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \mathcal{K}(\mathbf{p}) + \mathcal{V}(\mathbf{r}), \quad (1)$$

where

$$\mathcal{K}(\mathbf{p}) = \sum_{i=1}^N \mathbf{p}_i^2 / 2m_i \quad (2)$$

is the kinetic energy expressed in Cartesian coordinates as a Euclidean quadratic form of the momenta

$$\mathbf{p} := \{p_i = m_i \dot{r}_i, \quad i = 1, \dots, N\}, \quad (3)$$

and  $\mathcal{V} = \mathcal{V}(\mathbf{r})$  is a function of the  $3N$  Cartesian coordinates

$$\mathbf{r} := \{r_i, \quad i = 1, \dots, N\}. \quad (4)$$

$N$  is the number of point particles in the system, and we have put ourselves in dimension 3. The space of the coordinates is called configuration space, while the phase space  $\{\mathbf{r}, \mathbf{p}\}$  gives the space of the mechanical states of the system. To say that the system is subjected to  $f$  holonomic constraints is equivalent to saying that the motion has to evolve on a  $(3N - f)$ -dimensional configuration space, which results from imposing  $f$  geometrical conditions  $\sigma_\alpha(\mathbf{r}) = 0$ ,  $\alpha = 1, \dots, f$  at all times. These constraints can connect all the coordinates of the configuration space, in which case we call them global (Blue Moon [1,2]), or connect disjoint subgroups of the coordinates, which is, for example, the way in which molecular systems can sometimes be described [3], or else can be the conditions for orthonormality of single electron orbitals, as in the Car–Parrinello approach to ab initio molecular dynamics [4].

Global constraints can be used to bring the system in situations normally difficult to visit. In these cases, the constraints can act as a kind of Maxwell daemon. In the solution of the classical Statistical Mechanics of dynamical systems, the constraints confront two major problems. As we have seen before, the constraints are an essential ingredient in the definition of the dynamical system, therefore any acceptable algorithm introduced to solve the dynamics of such a system cannot propagate any error, as otherwise the statistical behavior of the ensemble in the presence of the constraints cannot be properly formulated. This last problem in principle is automatically solved for Hamiltonian systems by using generalized coordinates. However, especially for systems with many degrees of freedom, generalized coordinates are completely intractable, and one should be able to formulate properly the problem by using Cartesian coordinates in a standard way. As we will see, a family of algorithms avoiding the propagation of the errors have been introduced [3,5], while the proper formulation of the statistical ensemble is straightforward for the equilibrium case [6] but requires some more work for non-equilibrium, where the missing ingredient is the correct Liouville equation to use [7]. To get the proper Liouville equation one has to abandon the traditional Hamiltonian description, in which one had the constraint forces by using a Lagrange multiplier, and go straight to the non-Hamiltonian behavior of the equations of motion of the system in which the constraint forces are explicitly (analytically) solved. For this non-Hamiltonian equations the Euclidean nature of phase space is no more an invariant, therefore one has to find an invariant, non-Euclidean, measure to be associated with the phase space so that the statistical behavior of the system can be properly described by generalizing the Liouville equation [7]. As we will see, the non-equilibrium response of our constrained system to external perturbations can be derived directly [8], while the best known results of linear response theory, the so-called fluctuation-dissipation theorem, can be derived but requires some extra work [9]. The advised reader should be warned that in this review we have excluded the treatment of non-holonomic constraints, a very large family difficult to unify and in any event requiring special treatments [10].

In Section 2, we summarize the formalism needed to describe a system with constraints. In Section 3, we write down the general (SHAKE) equations to be solved to derive whatever family of numerical algorithms and briefly describe the two best known formulations: the one, adopted with the Verlet algorithm, usually referred to as SHAKE [11], not to be confused with the reference to the general equation to be solved, and the one modified to work with the velocity-Verlet algorithm, usually referred to as RATTLE [12]. To these two, we will briefly add a more recent alternative devised to give a parallel implementation of SHAKE [13]. In Section 4, we derive the equilibrium ensemble of Hamiltonian systems subjected to holonomic constraints [6]. Section 5 presents an effective approach to compute conditional averages by the use of holonomic constraints (Blue Moon) [1,2,14,15].

The possibility to compute conditional averages can be used in conjunction with non-equilibrium molecular dynamics techniques [16–20] to compute rate constants, hydrodynamical phenomena and, in general, relaxations from large fluctuations statistically produced by introducing suitable constraints. In Section 6, we formulate the non-Hamiltonian equations of motion for a constrained system, we derive from them an invariant measure for the phase space, the correct generalized Liouville equation and, again, as a way to see how all that works, the expressions for the equilibrium ensembles. Then, we start from the generalized Liouville equation to give a rigorous expression for the response to external perturbations of a constrained system [8] and we prove with some rigorous arguments that also the classical results of linear response theory can be shown to hold [9]. The paper is concluded by a short outlook in which we try to assess the state of the art in the treatment of the computational classical Statistical Mechanics for systems subjected to holonomic constraints.

## 2. Dynamics with Holonomic Constraints

Given the Lagrangian  $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}) = \mathcal{K}(\dot{\mathbf{r}}) - \mathcal{V}(\mathbf{r})$  of a dynamical system with  $N$  particles in dimension 3, subjected to  $f$  holonomic constraints

$$\sigma_\alpha(\mathbf{r}) = 0, \quad \alpha = 1, \dots, f, \tag{5}$$

the equations of motion (Lagrange equation of I type) are

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} - \frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = m_i \ddot{\mathbf{r}}_i + \frac{\partial \mathcal{V}(\mathbf{r})}{\partial \mathbf{r}_i} = - \sum_{\alpha=1}^f \lambda_\alpha(t) \frac{\partial \sigma_\alpha(\mathbf{r})}{\partial \mathbf{r}_i}, \quad i = 1, \dots, N, \tag{6}$$

where the  $\lambda_\alpha(t)$ ,  $\alpha = 1, \dots, f$  are the unknown Lagrangian multipliers to be determined by imposing that

$$\sigma_\alpha(\mathbf{r}(t)) = 0, \quad \forall t, \quad \alpha = 1, \dots, f. \tag{7}$$

Note that the multipliers  $\lambda_\alpha(t)$  together with all their derivatives of any order can be determined by taking successive time derivatives of the expressions  $\sigma_\alpha(\mathbf{r}(t)) = 0$ . In particular,

$$\dot{\sigma}_\alpha(t) = \dot{\sigma}_\alpha(\mathbf{r}(t), \dot{\mathbf{r}}(t)) = \sum_{i=1}^N \left( \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} \right) \sigma_\alpha(\mathbf{r}) = 0, \quad \alpha = 1, \dots, f \tag{8}$$

are evident conditions to be satisfied since holonomic constraints do not perform any mechanical work and therefore permitted velocities and constrained forces have to be orthogonal. Moreover,

$$\ddot{\sigma}_\alpha(t) = \left[ \sum_{i=1}^N \left( \ddot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} \right) + \sum_{i=1}^N \sum_{j=1}^N \left( \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} \right) \left( \dot{\mathbf{r}}_j \cdot \frac{\partial}{\partial \mathbf{r}_j} \right) \right] \sigma_\alpha(\mathbf{r}) = 0, \quad \alpha = 1, \dots, f. \tag{9}$$

Substituting in Equation (9) the equations of motion, Equation (6), and solving the resulting linear system for the  $\lambda$ s, we get

$$\lambda_\alpha = \sum_{\beta=1}^f \left[ - \frac{1}{m_i} \frac{\partial \mathcal{V}}{\partial \mathbf{r}_i} \cdot \left( \frac{\partial}{\partial \mathbf{r}_i} \right) \sigma_\beta + \sum_{i=1}^N \sum_{j=1}^N \left( \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} \right) \left( \dot{\mathbf{r}}_j \cdot \frac{\partial}{\partial \mathbf{r}_j} \right) \sigma_\beta \right] Z_{\beta,\alpha}^{-1}, \quad \alpha = 1, \dots, f, \tag{10}$$

where

$$Z_{\alpha,\beta}(\mathbf{r}) = \sum_{i=1}^N \frac{1}{m_i} \left( \frac{\partial \sigma_\alpha}{\partial \mathbf{r}_i} \right) \cdot \left( \frac{\partial \sigma_\beta}{\partial \mathbf{r}_i} \right), \quad \alpha = 1, \dots, f, \beta = 1, \dots, f. \tag{11}$$

The expressions resulting from introducing Equation (10) in Equation (6) could possibly provide fully explicit, no longer Hamiltonian, dynamics. For the moment, we will not be interested in such formulation because any approximate algorithm that make use of Equation (10) will necessarily propagate the errors in the constraint relations, Equation (5), with dramatic consequences on the stability of the model. Further derivatives with respect to time of the  $\sigma$ s will provide linear relationships for the higher order derivatives of the  $\lambda$ s, which could be needed in higher order algorithms.

By taking the standard Legendre transform on the Lagrangian  $\mathcal{L}$ ,

$$\mathcal{H} = \sum_{i=1}^N p_i \cdot \dot{r}_i - \mathcal{L}, \tag{12}$$

the same dynamics can be straightforwardly formulated in Hamiltonian terms,

$$\dot{r}_i = \frac{p_i}{m_i}, \tag{13}$$

$$\dot{p}_i = -\frac{\partial \mathcal{V}(\mathbf{r})}{\partial r_i} - \sum_{\alpha=1}^f \lambda_{\alpha}(t) \frac{\partial \sigma_{\alpha}(\mathbf{r})}{\partial r_i}, \quad i = 1, \dots, N, \tag{14}$$

involving the same treatment for the constraint forces. In the following, it will be useful for theoretical purposes to consider an equivalent representation of the Hamiltonian expressed in terms of

- (i) the  $f$  constraint relationships  $\sigma := \{\sigma_{\alpha}(\mathbf{r}), \alpha = 1, \dots, f\}$ , and
- (ii) the remaining  $(3N - f)$  generalized coordinates  $\mathbf{q} := \{q_{\nu}(\mathbf{r}), \nu = 1, \dots, 3N - f\}$ .

This change of coordinates is a point transformation of the configuration space and therefore generates a canonical transformation [10]. With this change of variables, the Lagrangian  $\mathcal{L}$  of our system generates the Lagrangian in the new coordinates given by

$$\mathcal{L}^*(\mathbf{q}, \dot{\mathbf{q}}, \sigma, \dot{\sigma}) = \mathcal{L}(\mathbf{r}(\mathbf{q}, \sigma), \dot{\mathbf{r}}(\mathbf{q}, \sigma, \dot{\mathbf{q}}, \dot{\sigma})). \tag{15}$$

Sometimes, it is useful to call collectively the variable  $\{\mathbf{q}, \sigma\} = \{\mathbf{u}\}$ . From the Lagrangian  $\mathcal{L}^*$ , we get

$$\mathbf{p}_u = \frac{\partial \mathcal{L}^*}{\partial \dot{\mathbf{u}}} \quad \left( \mathbf{p}_q = \frac{\partial \mathcal{L}^*}{\partial \dot{\mathbf{q}}}, \mathbf{p}_{\sigma} = \frac{\partial \mathcal{L}^*}{\partial \dot{\sigma}} \right) \tag{16}$$

and

$$\mathcal{H}^*(\mathbf{u}, \mathbf{p}_u) = \mathcal{K}(\mathbf{p}(\mathbf{u}, \mathbf{p}_u)) + \mathcal{V}(\mathbf{r}(\mathbf{u})) = \mathcal{K}^*(\mathbf{u}, \mathbf{p}_u) + \mathcal{V}^*(\mathbf{u}), \tag{17}$$

where the kinetic term is now

$$\mathcal{K}^* = \frac{1}{2} \mathbf{p}_u^T \cdot \mathbf{M}^{-1}(\mathbf{u}) \cdot \mathbf{p}_u \tag{18}$$

and

$$(\mathbf{M}(\mathbf{u}))_{\alpha, \beta} = \sum_{i=1}^N m_i \frac{\partial r_i}{\partial u_{\alpha}} \cdot \frac{\partial r_i}{\partial u_{\beta}}, \alpha = 1, \dots, 3N, \beta = 1, \dots, 3N \tag{19}$$

is the metric matrix associated with the new variables. Note that it is almost immediate to find for the inverse matrix  $\mathbf{M}^{-1}$  the explicit expression

$$(\mathbf{M}^{-1}(\mathbf{r}))_{\alpha, \beta} = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial u_{\alpha}}{\partial r_i} \cdot \frac{\partial u_{\beta}}{\partial r_i}, \alpha = 1, \dots, 3N, \beta = 1, \dots, 3N. \tag{20}$$

There is an intimate connection between the matrix  $M$ , the Jacobian matrix  $J = \frac{\partial r}{\partial u}$  and the Jacobian determinant  $J = \frac{\partial(u)}{\partial(r)} = |J|$  of the  $\{r\} \rightarrow \{u\}$  point transformation. Introducing the mass tensor  $(\mu)_{i,k} = m_i \delta_{i,k}$  for  $i, k = 1, \dots, 3N$  by means of the Kronecker delta, one can rewrite  $M$  in Equation (19) and, of course, its determinant  $|M|$ , as

$$M = \left( \frac{\partial r}{\partial u} \right)^T \cdot \mu \cdot \left( \frac{\partial r}{\partial u} \right) = J^T \cdot \mu \cdot J \quad \implies \quad |M| = |J| |\mu| |J| = |\mu| J^2. \tag{21}$$

In this representation, the constrained motion is generated by the Lagrangian  $\mathcal{L}^*(q, \dot{q}, \sigma = 0, \dot{\sigma} = 0) = \mathcal{L}_c(q, \dot{q})$  in the  $(3N - f)$ -dimensional space

$$\frac{\partial \mathcal{L}_c}{\partial q_\nu} - \frac{d}{dt} \frac{\partial \mathcal{L}_c}{\partial \dot{q}_\nu} = 0, \quad \nu = 1, \dots, 3N - f. \tag{22}$$

Note, however, that these equations are no longer in normal form.

The Hamiltonian formulation helps us to get back to an evolution expressed in normal form. Given that

$$p_u = \frac{\partial \mathcal{L}^*}{\partial \dot{u}} = M(u) \cdot \dot{u} \quad \text{i.e.,} \quad \dot{u} = M^{-1}(u) \cdot p_u, \tag{23}$$

we have

$$\dot{u} = \frac{\partial \mathcal{H}^*}{\partial p_u} = M^{-1}(u) \cdot p_u, \tag{24}$$

$$\dot{p}_u = - \frac{\partial \mathcal{H}^*}{\partial u}. \tag{25}$$

To proceed, it is useful to write the matrices  $M$  and  $M^{-1}$  in block form

$$M = \begin{pmatrix} A & B \\ B^T & \Gamma \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} \Delta & E \\ E^T & Z \end{pmatrix}, \tag{26}$$

where

$$A_{\nu,\eta}(q, \sigma) = \sum_{i=1}^N m_i \frac{\partial r_i}{\partial q_\nu} \cdot \frac{\partial r_i}{\partial q_\eta}, \quad \nu = 1, \dots, 3N - f, \quad \eta = 1, \dots, 3N - f, \tag{27}$$

$$B_{\nu,\alpha}(q, \sigma) = \sum_{i=1}^N m_i \frac{\partial r_i}{\partial q_\nu} \cdot \frac{\partial r_i}{\partial \sigma_\alpha}, \quad \nu = 1, \dots, 3N - f, \quad \alpha = 1, \dots, f, \tag{28}$$

$$\Gamma_{\alpha,\beta}(q, \sigma) = \sum_{i=1}^N m_i \frac{\partial r_i}{\partial \sigma_\alpha} \cdot \frac{\partial r_i}{\partial \sigma_\beta}, \quad \alpha = 1, \dots, f, \quad \beta = 1, \dots, f, \tag{29}$$

$$\Delta_{\nu,\eta}(r) = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial q_\nu}{\partial r_i} \cdot \frac{\partial q_\eta}{\partial r_i}, \quad \nu = 1, \dots, 3N - f, \quad \eta = 1, \dots, 3N - f, \tag{30}$$

$$E_{\nu,\alpha}(r) = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial q_\nu}{\partial r_i} \cdot \frac{\partial \sigma_\alpha}{\partial r_i}, \quad \nu = 1, \dots, 3N - f, \quad \alpha = 1, \dots, f, \tag{31}$$

with  $Z$  already defined in Equation (11), to derive a number of results, which we will use in the following. In particular, note that the block matrices defined above are not independent from each

other. A first set of useful relations can be derived by expanding the expressions for the identity  $\mathbf{1} = M^{-1}M = MM^{-1}$

$$M^{-1}M = \begin{pmatrix} \Delta A + EB^T = \mathbf{1} & \Delta B + E\Gamma = \mathbf{0} \\ E^T A + ZB^T = \mathbf{0} & E^T B + Z\Gamma = \mathbf{1} \end{pmatrix} = \begin{pmatrix} A\Delta + BE^T = \mathbf{1} & AE + BZ = \mathbf{0} \\ B^T\Delta + \Gamma E^T = \mathbf{0} & B^T E + \Gamma Z = \mathbf{1} \end{pmatrix} = MM^{-1}. \tag{32}$$

Another useful relation is the one that, in a different language, is known as Fixman’s Theorem [21]. It relates the determinants of the matrices  $A$  and  $Z$

$$|A| = |M||Z|. \tag{33}$$

Equation (33) can be derived directly observing that (use Equation (32))

$$(i) \quad \begin{pmatrix} A & \mathbf{0} \\ B^T & \mathbf{1} \end{pmatrix} = M \cdot \left[ M^{-1} \cdot \begin{pmatrix} A & \mathbf{0} \\ B^T & \mathbf{1} \end{pmatrix} \right] = M \cdot \begin{pmatrix} \mathbf{1} & E \\ \mathbf{0} & Z \end{pmatrix}, \tag{34}$$

and that

$$(ii) \quad \left| \begin{pmatrix} A & \mathbf{0} \\ B^T & \mathbf{1} \end{pmatrix} \right| = |A|, \quad \left| \begin{pmatrix} \mathbf{1} & E \\ \mathbf{0} & Z \end{pmatrix} \right| = |Z|. \tag{35}$$

By putting Equation (33) together with Equation (21), we obtain for the determinant of the block matrix  $A$ , the interesting expression

$$|A|^{\frac{1}{2}} = (|M||Z|)^{\frac{1}{2}} = J |\mu|^{\frac{1}{2}} |Z|^{\frac{1}{2}} \tag{36}$$

that will be useful later on.

Going back to the constraint relations, the conditions  $\dot{\sigma} = 0$  can now be written as

$$\dot{\sigma} = E^T p_q + Z p_\sigma = 0, \tag{37}$$

giving the non-zero values of the conjugated momenta  $\tilde{p}_\sigma$  when the constraints are imposed,

$$\tilde{p}_\sigma = -Z^{-1}(q, \sigma = 0) E^T(q, \sigma = 0) p_q = -\tilde{Z}^{-1} \tilde{E}^T p_q, \tag{38}$$

where  $\tilde{Z}$  and  $\tilde{E}$  are implicitly defined in Equation (38). In these conditions, the Hamiltonian of the constrained motion can be evaluated explicitly on the hypersurface  $\sigma = 0, p_\sigma = \tilde{p}_\sigma$  to obtain

$$\mathcal{H}^*(q, \sigma = 0, p_q, p_\sigma = \tilde{p}_\sigma) = \mathcal{K}^*(q, \sigma = 0, p_q, p_\sigma = \tilde{p}_\sigma) + \mathcal{V}^*(q, \sigma = 0) \tag{39}$$

$$= \frac{1}{2} (p_q^T \tilde{p}_\sigma^T) \begin{pmatrix} \tilde{\Delta} & \tilde{E} \\ \tilde{E}^T & \tilde{Z} \end{pmatrix} \begin{pmatrix} p_q \\ \tilde{p}_\sigma \end{pmatrix} + \mathcal{V}_c(q) \tag{40}$$

$$= \frac{1}{2} p_q^T (\tilde{\Delta} - \tilde{E} \tilde{Z}^{-1} \tilde{E}^T) p_q + \mathcal{V}_c(q) \tag{41}$$

$$= \frac{1}{2} p_q^T \tilde{A}^{-1}(q) p_q + \mathcal{V}_c(q) \equiv \mathcal{H}_c(q, p_q), \tag{42}$$

where we first used  $\tilde{p}_\sigma = -\tilde{Z}^{-1} \tilde{E}^T p_q$  to go from Equation (40) to Equation (41) and, then, from Equation (41) to Equation (42), we have used the relations from Equation (32),  $A\Delta + BE^T = \mathbf{1}$  and  $AE + BZ = \mathbf{0}$ . The Hamiltonian  $\mathcal{H}_c$  generates the equations of motion in the  $(3N - f)$ -dimensional space in normal form

$$\dot{q} = \frac{\partial \mathcal{H}_c}{\partial p_q}, \tag{43}$$

$$\dot{p}_q = -\frac{\partial \mathcal{H}_c}{\partial q}. \tag{44}$$

### 3. SHAKE, Integrating the Equations of Motion

The numerical integration of the equations of motion (6) requires discretizing the time and to provide a suitable algorithm of given precision. Generally, for evident reasons, one avoids the use of algorithms requiring more than the computation at each step of the forces  $F_i = -\frac{\partial \mathcal{V}}{\partial r_i}$  avoiding successive derivatives, e.g.,  $\dot{F}_i = -\left(\sum_j \dot{r}_j \cdot \frac{\partial}{\partial r_j}\right) \left(\frac{\partial \mathcal{V}}{\partial r_i}\right)$ , etc. For illustrative purposes, we will limit ourselves to write down the integration of the Lagrangian equations of motion using the Verlet algorithm and of the Hamiltonian equations using the velocity-Verlet algorithm.

#### 3.1. Verlet Algorithm

The celebrated (1967) Verlet algorithm is easily obtained by writing down and summing up the forward and backward Taylor expansions of each coordinate truncated to the fourth order. Calling  $x$  a generic variable in the configuration space set  $\mathbf{r}$  (and  $\partial_x = \frac{\partial}{\partial x}$ ), it reads

$$x(t+h) = -x(t-h) + 2x(t) + h^2\ddot{x}(t) + O(h^4), \tag{45}$$

where  $t$  is the running time and  $h$  is the integration step resulting from time discretization. The velocity with this algorithm is computed by subtracting the same forward and backward Taylor expansions. We get, with one timestep of delay,

$$\dot{x}(t) = \frac{x(t+h) - x(t-h)}{2h} - h^2\frac{\ddot{x}(t)}{3} + O(h^4), \tag{46}$$

where we have written explicitly the error of order  $O(h^2)$  for further use. Notice that the velocities, which carry a larger error, do not enter in the computation of the trajectory, which remains precise to the order three, and, as it has been shown, has many other remarkable features that can be summarized by saying that this algorithm is symplectic [22,23]. In presence of holonomic constraints, the acceleration of any coordinate  $x$  can be decomposed in the two contributions,  $F = -\partial_x \mathcal{V}$  coming from the interaction potential of the model and  $G = -\sum_{\alpha=1}^f \Lambda_\alpha (\partial_x \sigma_\alpha) \equiv -\Lambda \cdot (\partial_x \sigma)$  the constraint force, where  $\Lambda$  is a set of parameters to be determined, so that

$$\ddot{x}(t) = \frac{1}{m} [F(t) + G(t)]. \tag{47}$$

Substituting Equation (47) in the algorithm Equation (45), we have

$$x(t+h) = \bar{x}(t+h) + \frac{h^2}{m} G(t) + O(h^4), \tag{48}$$

where  $\bar{x}(t+h)$ , the provisional value of the coordinate at time  $t+h$ , is the position the coordinate would take in the absence of constraints. Now, and this is the essential conceptual content of the whole family of SHAKE algorithms, we determine the set of the  $\Lambda$  parameters by imposing and solving the set of algebraic equations

$$\sigma_\alpha \left( \left\{ \bar{x}(t+h) - \frac{h^2}{m} \Lambda \cdot (\partial_x \sigma) \right\} \right) = 0, \quad \alpha = 1, \dots, f. \tag{49}$$

Since we have  $f$   $\Lambda$  values and  $f$  constraint relationships, the system of algebraic, generally not linear, equations is well posed. The values of the  $\Lambda$ s solving these equations are in general different from the values of the  $\lambda(t)$  obtainable from Equations (10). However, the difference cannot be of greater order than the one involved in the algorithmic error. Therefore, the values of the coordinates at time  $t + h$  will entail an error equivalent to the one produced by the blind application of the Verlet algorithm. However, now, the constraint relationships will be satisfied exactly at every timestep and the dynamics of the system will not disrupt the model.

Many different ways have been proposed to solve the system of Equations (46), see e.g., [13,24–27]. The original, and still commonly used, goes back to Berendsen [3], who called it, again, SHAKE. It proceeds by satisfying one constraint at a time, iterating constraint relationship by constraint relationship until convergence. The technical details have been worked out, apart from the original paper, more pedagogically in [11]. Leimkuhler [28] has demonstrated that the resulting numerical procedure maintains the time reversal invariance and the symplectic character of the algorithm. In the referred to, original, implementation, the algorithm is inherently serial and cannot be easily parallelized. Practical parallelizations are either approximate or the algorithms are specifically tailored to the problems at hand. An interesting general parallel solution has been worked out by Weinbach and Elber [13]. They take advantage of the fact that the essential step in solving the SHAKE equation

$$\sigma_\alpha(\bar{x}(t+h), \Lambda) = 0, \quad \alpha = 1, \dots, f \tag{50}$$

can be recast as the solution of a sparse linear problem of the type  $\mathbf{A}\mathbf{y} = \mathbf{b}$  with  $\mathbf{y}$  the vector of unknowns. Constructing a suitable positive definite matrix, they solve the SHAKE equation using (parallel) conjugate gradient minimization of the quadratic form  $\frac{1}{2}\Lambda^T \mathbf{A}\Lambda - \Lambda^T \sigma$  in place of the standard iterative process (inherently serial).

### 3.2. Velocity-Verlet Algorithm

The Verlet algorithm can be easily recast in an algebraically equivalent form that, when applied in the correct order, produces both the positions and the velocities at the same time. Using the same symbols, let us first rewrite Equation (45) by replacing  $-x(t-h)$  with its value extracted from Equation (46)

$$-x(t-h) = 2h\dot{x}(t) - x(t+h) + \frac{2h^3}{3}\ddot{x}(t+h) + O(h^5), \tag{51}$$

where we have written explicitly for further use the expression for the error  $O(h^3)$ , to be normally rejected, in order to obtain the first equation of the velocity-Verlet algorithm, which expresses the position  $x$  at time  $(t+h)$

$$x(t+h) = x(t) + h\dot{x}(t) + \frac{h^2}{2}\ddot{x}(t) + h^3\frac{\ddot{x}(t)}{3} + O(h^5) \tag{52}$$

with an error of order  $O(h^3)$ , here retained in its explicit form. Next, we write the velocity  $\dot{x}(t+h)$  from Verlet (46) taken at time  $(t+h)$  and eliminate the position  $x(t+2h)$  using, again, Verlet (45) taken to go from time  $(t+h)$  to time  $(t+2h)$ ,

$$\dot{x}(t+h) = [x(t+2h) - x(t)] / (2h) - \frac{h^2}{3} \ddot{x}(t+h) + O(h^4) \tag{53}$$

$$= \left[ 2x(t+h) - x(t) + h^2 \ddot{x}(t+h) + O(h^4) - x(t) \right] / (2h) - \frac{h^2}{3} \ddot{x}(t+h) + O(h^4) \tag{54}$$

$$= \frac{x(t+h) + (2x(t) - x(t-h) + h^2 \ddot{x}(t) + O(h^4)) - x(t) + h^2 \ddot{x}(t+h) + O(h^4) - x(t)}{2h} - \frac{h^2}{3} \ddot{x}(t+h) + O(h^4), \tag{55}$$

where we made again use of Verlet (45) to expand one of the two  $x(t+h)$  contributions. Grouping terms and simplifying, we can write

$$\dot{x}(t+h) = \frac{x(t+h) - x(t-h)}{2h} + \frac{h}{2} [\ddot{x}(t) + \ddot{x}(t+h)] + O(h^3) - \frac{h^2}{3} \ddot{x}(t+h), \tag{56}$$

$$= \dot{x}(t) + \frac{h}{2} [\ddot{x}(t) + \ddot{x}(t+h)] + \frac{h^2}{3} [\ddot{x}(t) - \ddot{x}(t+h)] + O(h^3), \tag{57}$$

where we got rid of the fraction

$$\frac{x(t+h) - x(t-h)}{2h} = \dot{x}(t) + h^2 \frac{\ddot{x}(t)}{3} + O(h^4), \tag{58}$$

using Verlet (46). Finally, by expanding with Taylor the third derivative term at time  $(t+h)$ ,  $\ddot{x}(t+h) = \ddot{x}(t) + h \ddot{x}'(t) + O(h^2)$ , we observe that the two terms  $\propto h^2$  cancel each other leaving a term in  $h^3$ . Finally, we arrive at the second equation of the velocity-Verlet algorithm

$$\dot{x}(t+h) = \dot{x}(t) + \frac{h}{2} [\ddot{x}(t) + \ddot{x}(t+h)] + O(h^3), \tag{59}$$

which expresses the velocity  $\dot{x}$  at time  $(t+h)$ , again with an error of order  $O(h^3)$ .

Substituting Equation (47) into Equation (52), we have, analogously to the previous section,

$$x(t+h) = \tilde{x}(t+h) + \frac{h^2}{2m} G'(t) + O(h^3), \tag{60}$$

where  $\tilde{x}(t+h)$ , the provisional value of the coordinate at time  $t+h$ , is the position that the coordinate would take in the absence of constraints using Equation (52), and  $G' = -\Lambda' \cdot (\partial_x \sigma)$  is the constraint force and the parameters  $\Lambda'$  are determined by imposing and solving the set of equations

$$\sigma_\alpha \left( \left\{ \tilde{x}(t+h) - \frac{h^2}{2m} \Lambda' \cdot (\partial_x \sigma) \right\} \right) = 0, \quad \alpha = 1, \dots, f. \tag{61}$$

Again, we have an algebraic system with  $f$  constraint relationships  $\sigma$  and  $f$  unknown values  $\Lambda'$ . The problem is well posed and the solution can be retrieved exactly along the same lines as before, by using the iterative SHAKE algorithm. Of course, we will have a different set of parameters,  $\Lambda' \neq \lambda$ , but the new positions  $x(t+h)$  will satisfy the constraints (5) exactly at the time  $t+h$ . To calculate the new velocities, the above procedure must be repeated by substituting Equation (47) in Equation (59), now at time  $t+h$ ,

$$\dot{x}(t+h) = \dot{\tilde{x}}(t+h) + \frac{h}{2m} G''(t+h) + O(h^3), \tag{62}$$

where  $G'' = -\Lambda'' \cdot (\partial_x \sigma)$  and

$$\tilde{x}(t+h) = \dot{x}(t) + \frac{h}{2} [F(t) + F(t+h)] + \frac{h}{2m} G'(t), \tag{63}$$

the provisional velocity at time  $t+h$ , i.e., the value the velocity would take in the absence of constraints at time  $(t+h)$ . Note that, at this stage, the  $\Lambda'$  and, therefore, the constraint force  $G'(t)$  at time  $t$  are already computed and therefore included in  $\tilde{x}(t+h)$ , while one needs to determine the yet unknown parameters  $\Lambda''$  by imposing and solving the set of equations

$$\dot{\sigma}_\alpha \left( \{x(t+h), \tilde{x}(t+h) - \frac{h}{2m} \Lambda'' \cdot (\partial_x \sigma)(t+h)\} \right) = 0, \quad \alpha = 1, \dots, f. \tag{64}$$

Once more, we have an algebraic system with  $f$  constraint relationships,  $\dot{\sigma}$ , and  $f$  unknown values,  $\Lambda''$ . The problem is well posed and the solution can be retrieved by an iterative SHAKE-like procedure, i.e., proceeding by satisfying one constraint relation at a time. The whole procedure of imposing constraints within the velocity-Verlet scheme is known by a different name, the RATTLE algorithm [12], although it is indeed nothing else than the same SHAKE procedure applied twice, once for positions and once for velocities, to two different sets of equations. The main difference with the original SHAKE algorithm [3] lies in the fact that the velocities calculated using Equation (62) are at each time exactly tangent to the constraint hypersurface  $\sigma = 0$ , while the velocities calculated, usually, simply using Equation (46) in the original SHAKE algorithm are tangent only within the algorithm accuracy ( $O(h^2)$ ). This extra precision does not come for free, but at the cost of doubling the effort in calculating the unknown  $\Lambda$  parameters. As long as the velocities in the Verlet algorithm do not enter directly into the numerical integration of the positions, such difference can be safely ignored; however, if desired, nothing would impede applying SHAKE in the same spirit, and with similar costs, to correct the velocities from Equation (46).

#### 4. Equilibrium Statistical Mechanics in the Hamiltonian Formulation

The expression of the statistical equilibrium ensemble in Cartesian coordinates of a system subjected to holonomic constraints is not smooth but singular since the probability density defined in a  $6N$ -dimensional phase space is associated with a mechanical system whose motion takes place in a  $(6N - 2f)$ -dimensional subspace, i.e., the intersection of the  $2f$  hypersurfaces  $\sigma(\mathbf{r}) = 0$  and  $\dot{\sigma}(\mathbf{r}, \mathbf{p}) = 0$ .

On the contrary, it is immediate to write down the (microcanonical) probability density in the reduced phase space of the  $2(3N - f)$  generalized coordinates  $\mathbf{q}, \mathbf{p}_q$  using the Hamiltonian  $\mathcal{H}_c$  in Equation (42).

Let  $\hat{\mathcal{O}}(\mathbf{r}, \mathbf{p})$  be a dynamical variable defined using Cartesian coordinates and  $\hat{\mathcal{O}}_c(\mathbf{q}, \mathbf{p}_q) = \hat{\mathcal{O}}(\mathbf{r}(\mathbf{q}, \sigma = 0), \mathbf{p}(\mathbf{q}, \sigma = 0, \mathbf{p}_q, \mathbf{p}_\sigma = \tilde{\mathbf{p}}_\sigma))$  the equivalent variable expressed using generalized coordinates, restricted to the constrained hypersurface. The familiar microcanonical average in generalized coordinates reads

$$\langle \hat{\mathcal{O}} \rangle_{NVE} = \frac{1}{N! \Omega(N, V, E)} \int d\mathbf{q} d\mathbf{p}_q \hat{\mathcal{O}}_c(\mathbf{q}, \mathbf{p}_q) \delta(\mathcal{H}_c(\mathbf{q}, \mathbf{p}_q) - E), \tag{65}$$

where

$$\Omega(N, V, E) = \frac{1}{N!} \int d\mathbf{q} d\mathbf{p}_q \delta(\mathcal{H}_c(\mathbf{q}, \mathbf{p}_q) - E). \tag{66}$$

We will now transform it into the equivalent integral in Cartesian coordinates by making use of the canonical transformation that connect the “generalized” phase space variables  $(\mathbf{u}, \mathbf{p}_u)$  introduced

in Section 2 to the “Cartesian” phase space variables  $(\mathbf{r}, \mathbf{p})$ . We first remark that, on the  $(6N - 2f)$  phase space hypersurface, one has

$$d\mathbf{q} d\mathbf{p}_q = dq d\sigma \delta(\sigma) d\mathbf{p}_q d\mathbf{p}_\sigma \delta(\mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma) = du d\mathbf{p}_u \delta(\sigma) \delta(\mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma), \tag{67}$$

where, for the product of the first  $f$  delta functions, we have used the shortcut notation  $\delta(\sigma) = \prod_\alpha \delta(\sigma_\alpha)$ , and, equivalently, for the last term  $\delta(\mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma) = \prod_\alpha \delta((\mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma)_\alpha)$ .

A more convenient expression for this product of delta functions can be derived by noting that by multiplying Equation (37) by  $\mathbf{Z}^{-1}$ , one obtains:

$$\mathbf{Z}^{-1} \dot{\sigma} = \mathbf{p}_\sigma + \mathbf{Z}^{-1} \mathbf{E}^T \mathbf{p}_q = \mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma. \tag{68}$$

Finally, using the facts that the Jacobian associated with a canonical transformation generated by a point transformation in the coordinates preserves the phase space volume, i.e.,  $\frac{\partial(\mathbf{u}, \mathbf{p}_u)}{\partial(\mathbf{r}, \mathbf{p})} = 1$ , and Equation (42), we can write for the microcanonical average (65)

$$\langle \hat{\mathcal{O}} \rangle_{NVE} = \frac{1}{N! \Omega(N, V, E)} \int du d\mathbf{p}_u \hat{\mathcal{O}}^*(\mathbf{u}, \mathbf{p}_u) \delta[\mathcal{H}^*(\mathbf{u}, \mathbf{p}_u) - E] \delta(\sigma) \delta(\mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma) \tag{69}$$

$$= \frac{1}{N! \Omega(N, V, E)} \int d\mathbf{r} d\mathbf{p} \hat{\mathcal{O}}(\mathbf{r}, \mathbf{p}) \delta[\mathcal{H}(\mathbf{r}, \mathbf{p}) - E] \delta(\sigma(\mathbf{r})) \delta(\mathbf{Z}^{-1} \dot{\sigma}(\mathbf{r}, \mathbf{p})), \tag{70}$$

where now,

$$\begin{aligned} \Omega(N, V, E) &= \frac{1}{N!} \int d\mathbf{r} d\mathbf{p} \delta[\mathcal{H}(\mathbf{r}, \mathbf{p}) - E] \delta(\sigma(\mathbf{r})) \delta(\mathbf{Z}^{-1} \dot{\sigma}(\mathbf{r}, \mathbf{p})) \\ &= \frac{1}{N!} \int d\mathbf{r} d\mathbf{p} |\mathbf{Z}(\mathbf{r})| \delta(\dot{\sigma}(\mathbf{r}, \mathbf{p})) \delta(\sigma(\mathbf{r})) \delta[\mathcal{H}(\mathbf{r}, \mathbf{p}) - E], \end{aligned} \tag{71}$$

with  $|\mathbf{Z}|$  the modulus of the determinant of the matrix  $\mathbf{Z}$ . From Equation (69), it follows directly the expression for the probability density in the microcanonical equilibrium ensemble in Cartesian coordinates

$$\mathcal{P}^{(micro)}(\mathbf{r}, \mathbf{p}) = \frac{\delta[\mathcal{H}(\mathbf{r}, \mathbf{p}) - E] |\mathbf{Z}(\mathbf{r})| \delta(\dot{\sigma}(\mathbf{r}, \mathbf{p})) \delta(\sigma(\mathbf{r}))}{N! \Omega(N, V, E)} \tag{72}$$

and similar for other equilibrium ensembles. In particular for the Canonical ensemble, where classically momenta and coordinates are explicitly independent, the probability density will result in being

$$\mathcal{P}(\mathbf{r}, \mathbf{p}) = \frac{\exp\{-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})\} |\mathbf{Z}(\mathbf{r})| \delta(\dot{\sigma}(\mathbf{r}, \mathbf{p})) \delta(\sigma(\mathbf{r}))}{\int d\mathbf{r} d\mathbf{p} \exp\{-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})\} |\mathbf{Z}(\mathbf{r})| \delta(\dot{\sigma}(\mathbf{r}, \mathbf{p})) \delta(\sigma(\mathbf{r}))}. \tag{73}$$

For theoretical purposes, as we will see in the following, it is very useful to write the ensemble in terms of the marginal configurational probability density  $\mathcal{P}_M(\mathbf{r})$  and the conditional probability density of the momenta  $\mathcal{P}_C(\mathbf{p}|\mathbf{r})$ :

$$\mathcal{P}(\mathbf{r}, \mathbf{p}) = \mathcal{P}_M(\mathbf{r}) \mathcal{P}_C(\mathbf{p}|\mathbf{r}), \tag{74}$$

where

$$\mathcal{P}_M(\mathbf{r}) d\mathbf{r} = \left( \int d\mathbf{p} \mathcal{P}(\mathbf{r}, \mathbf{p}) \right) d\mathbf{r} = \frac{\int d\mathbf{p} e^{-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})} \delta(\sigma(\mathbf{r})) \delta(\mathbf{Z}^{-1} \dot{\sigma}(\mathbf{r}, \mathbf{p}))}{\int d\mathbf{r} d\mathbf{p} e^{-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})} \delta(\sigma(\mathbf{r})) \delta(\mathbf{Z}^{-1} \dot{\sigma}(\mathbf{r}, \mathbf{p}))} d\mathbf{r}. \tag{75}$$

Explicitly,  $\mathcal{P}_M$  can be computed by first integrating out the delta functions for  $\dot{\sigma}$  in Equation (75), which amounts, after a change of variables, to the substitutions  $\mathbf{p}_\sigma = \tilde{\mathbf{p}}_\sigma$ , and then by executing the

Gaussian integrals involved in the canonical ensemble. Assuming that the standard result of the integration of a unidimensional Gaussian integral is known,

$$\int_{-\infty}^{+\infty} dx e^{-\frac{1}{2}x^2/a} = \sqrt{2\pi a}, \tag{76}$$

and, for a multidimensional Gaussian integral, diagonalizing the quadratic form and using the invariance of the determinant under unitary transformations, one derives immediately for the integral that, in  $n$  dimensions,

$$\int \dots \int d\mathbf{x} e^{-\frac{1}{2}(\mathbf{x}^T \cdot \mathbf{A}^{-1} \cdot \mathbf{x})} = (2\pi)^{\frac{n}{2}} |\mathbf{A}|^{\frac{1}{2}}. \tag{77}$$

Focusing on the numerator in Equation (75), making a change of variables in the integral by using  $d\mathbf{p} = J^{-1}d\mathbf{p}_u = J^{-1}d\mathbf{p}_q d\mathbf{p}_\sigma$ , one has

$$e^{-\beta\mathcal{V}}\delta(\boldsymbol{\sigma})d\mathbf{r} J^{-1} \int d\mathbf{p}_q d\mathbf{p}_\sigma e^{-\frac{\beta}{2}[\mathbf{p}_u^T \cdot \mathbf{M}^{-1} \cdot \mathbf{p}_u]}\delta(\mathbf{p}_\sigma - \tilde{\mathbf{p}}_\sigma) = e^{-\beta\mathcal{V}}\delta(\boldsymbol{\sigma})d\mathbf{r} J^{-1} \int d\mathbf{p}_q e^{-\frac{\beta}{2}[\mathbf{p}_q^T \cdot \mathbf{A}^{-1} \cdot \mathbf{p}_q]} \tag{78}$$

$$= e^{-\beta\mathcal{V}}\delta(\boldsymbol{\sigma})d\mathbf{r} \left(\frac{2\pi}{\beta}\right)^{\frac{3N-f}{2}} J^{-1}|\mathbf{A}|^{\frac{1}{2}} = \left(\frac{2\pi}{\beta}\right)^{\frac{3N-f}{2}} |\boldsymbol{\mu}|^{\frac{1}{2}} e^{-\beta\mathcal{V}}\delta(\boldsymbol{\sigma})|\mathbf{Z}|^{\frac{1}{2}}d\mathbf{r}. \tag{79}$$

To obtain Equation (79), we have proceeded as follows. First, we substitute the kinetic Hamiltonian term in the exponential with its “unconstrained” expression in Equation (40) and integrate over  $d\mathbf{p}_\sigma$  using Equation (42). Now, we perform the remaining multidimensional Gaussian integral over the remaining momenta  $\mathbf{p}_q$  by using Equation (77) and use Equation (36) to arrive at the result in Equation (79). Following the same procedure for the denominator and simplifying the constants in Equation (79), we finally get for the normalized marginal probability density

$$\mathcal{P}_M(\mathbf{r}) = \frac{e^{-\beta\mathcal{V}(\mathbf{r})}\delta(\boldsymbol{\sigma})|\mathbf{Z}|^{\frac{1}{2}}}{\int d\mathbf{r} e^{-\beta\mathcal{V}(\mathbf{r})}\delta(\boldsymbol{\sigma})|\mathbf{Z}|^{\frac{1}{2}}}. \tag{80}$$

Equation (80) tells us that the marginal probability density in configuration space, in the presence of constraints, is not simply  $\propto \exp(-\beta\mathcal{V})\delta(\boldsymbol{\sigma})$  but contains the biasing term  $|\mathbf{Z}(\mathbf{r})|^{\frac{1}{2}}$  coming from the limitations in momentum space induced by the constraints.

The conditional probability density in momentum space is given by

$$\mathcal{P}_C(\mathbf{p}|\mathbf{r}) = \frac{\mathcal{P}(\mathbf{r}, \mathbf{p})}{\mathcal{P}_M(\mathbf{r})} = \frac{e^{-\frac{\beta}{2}[\mathbf{p}^T \cdot \boldsymbol{\mu}^{-1} \cdot \mathbf{p}] - \beta\mathcal{V}}\delta(\boldsymbol{\sigma})\delta(\mathbf{Z}^{-1}\dot{\boldsymbol{\sigma}}(\mathbf{r}, \mathbf{p}))}{\int d\mathbf{p} e^{-\frac{\beta}{2}[\mathbf{p}^T \cdot \boldsymbol{\mu}^{-1} \cdot \mathbf{p}] - \beta\mathcal{V}}\delta(\boldsymbol{\sigma})\delta(\mathbf{Z}^{-1}\dot{\boldsymbol{\sigma}}(\mathbf{r}, \mathbf{p}))} \tag{81}$$

$$= \frac{e^{-\frac{\beta}{2}[\mathbf{p}^T \cdot \boldsymbol{\mu}^{-1} \cdot \mathbf{p}]}\delta(\mathbf{Z}^{-1}\dot{\boldsymbol{\sigma}}(\mathbf{r}, \mathbf{p}))}{(2\pi/\beta)^{\frac{3N-f}{2}} |\boldsymbol{\mu}|^{\frac{1}{2}} |\mathbf{Z}|^{\frac{1}{2}}} \tag{82}$$

$$= (2\pi/\beta)^{-\frac{3N-f}{2}} |\boldsymbol{\mu}|^{-\frac{1}{2}} e^{-\frac{\beta}{2}[\mathbf{p}^T \cdot \boldsymbol{\mu}^{-1} \cdot \mathbf{p}]}\delta(\mathbf{Z}^{-1}\dot{\boldsymbol{\sigma}}(\mathbf{r}, \mathbf{p})), \tag{83}$$

where to get the first equality we referred to Equation (80); for the next step, the result implicit in Equation (79); and, finally, then Equation (83). The configuration dependent factor  $|\mathbf{Z}(\mathbf{r})|^{\frac{1}{2}}$  in Equation (83) indicates that, when there are constraints, positions and momenta are no longer independent. In particular, the distribution of momenta becomes no more simply Maxwellian.

### 5. Rare Events and Blue Moon Ensemble

In the statistical mechanical treatment of macroscopic phenomena, one is interested in computing the properties of interest by identifying suitable observables, i.e., function of phase space,  $\hat{\xi}(\mathbf{r}, \mathbf{p})$ ,

although here, and for a while, we will focus on observables depending only on the configuration space and obtaining their macroscopic counterpart by taking an ensemble average (to be definite, let us choose to work with the canonical ensemble) of it,

$$\xi = \langle \hat{\xi} \rangle = \frac{1}{Q} \int dr d\mathbf{p} e^{-\beta \mathcal{H}(r, \mathbf{p})} \hat{\xi}(\mathbf{r}), \quad Q = \int dr d\mathbf{p} e^{-\beta \mathcal{H}(r, \mathbf{p})}. \tag{84}$$

More generally, given one observable, it can be instructive to compute the marginal probability density associated with it in the ensemble

$$\mathcal{P}_{\xi}(\xi') = \frac{1}{Q} \int dr d\mathbf{p} e^{-\beta \mathcal{H}(r, \mathbf{p})} \delta(\hat{\xi}(\mathbf{r}) - \xi') \equiv \langle \delta(\hat{\xi} - \xi') \rangle. \tag{85}$$

Macroscopically speaking, this probability density has a profound meaning since it can be associated, via the definition of the (Landau) free energy

$$\mathcal{W}_{\xi}(\xi') = -k_B T \ln \mathcal{P}_{\xi}(\xi') \tag{86}$$

to the reversible work needed to bring the physical system from a reference state to the value  $\xi = \xi'$ . This fact can be easily seen by taking the derivative with respect to  $\xi'$  of Equation (86)

$$\frac{d\mathcal{W}_{\xi}(\xi')}{d\xi'} = -k_B T \frac{d}{d\xi'} \ln \mathcal{P}_{\xi}(\xi') - \frac{k_B T}{\mathcal{P}_{\xi}(\xi')} \frac{d}{d\xi'} \mathcal{P}_{\xi}(\xi') \tag{87}$$

$$= -\frac{k_B T}{\mathcal{P}_{\xi}(\xi')} \int dr d\mathbf{p} \frac{e^{-\beta \mathcal{H}(r, \mathbf{p})}}{Q} \frac{d}{d\xi'} \delta(\hat{\xi}(\mathbf{r}) - \xi'). \tag{88}$$

In the same spirit of Section 2, we introduce the canonical transformation from the Cartesian coordinates  $\{r, \mathbf{p}\}$  to the generalized coordinates  $\{u, \mathbf{p}_u\}$ , where  $u = (q, \xi)$  with the set  $q$  suitably chosen. Now, using that  $\frac{d}{d\xi'} = -\left(\frac{\partial}{\partial \xi}\right)$ , integrating by parts, we arrive at

$$\frac{d\mathcal{W}_{\xi}(\xi')}{d\xi'} = \frac{\int du d\mathbf{p}_u \left(\frac{\partial \mathcal{H}^*}{\partial \xi}\right) e^{-\beta \mathcal{H}^*} \delta(\xi - \xi')}{\int du d\mathbf{p}_u e^{-\beta \mathcal{H}^*} \delta(\xi - \xi')} = \frac{\left\langle \left(\frac{\partial \mathcal{H}^*}{\partial \xi}\right) \delta(\hat{\xi} - \xi') \right\rangle}{\langle \delta(\hat{\xi} - \xi') \rangle}, \tag{89}$$

where  $\mathcal{H}^*$  is the Hamiltonian expressed in the generalized  $u$  coordinates, see Equation (17). In generalized coordinates, the kinetic term  $\mathcal{K}^*$  in the Hamiltonian (see Equation (18)) gives a non-zero contribution to the derivative, which is nothing but a geometrical correction that ultimately involves (see Appendix A) the Jacobian  $J = \frac{\partial(u)}{\partial(r)}$  of the coordinate transformation,

$$\frac{d\mathcal{W}_{\xi}(\xi')}{d\xi'} = \left\langle \left[ \left(\frac{\partial \mathcal{V}^*}{\partial \xi}\right) - k_B T \left(\frac{\partial \ln J}{\partial \xi}\right) \right] \delta(\hat{\xi} - \xi') \right\rangle / \langle \delta(\hat{\xi} - \xi') \rangle. \tag{90}$$

From Equation (90), we see that the derivative of  $\mathcal{W}_{\xi}$  is a conditional average, at a given value of the observable, of the generalized force acting on the system, i.e., typically a thermodynamic force. The evaluation of this expression as given in Equation (90) requires constructing explicitly the set of generalized coordinate  $u$ , something usually very cumbersome, and needs an ad hoc derivation in each particular case. As a matter of fact, it is possible to circumvent this technical difficulty [15] and derive expressions directly in terms of the function  $\xi(\mathbf{r})$  and its derivatives with respect to the Cartesian coordinates  $r$ . The “work” associated with this force is what we identify with reversible work. By thermodynamic integration, we can get the reversible work relative to a reference state and by exponentiation the probability density associated with the random variable  $\hat{\xi}$ .

In standard conditions, when  $\hat{\xi}$  is a unimodal random variable, the sampling of its probability density is an easy matter that can be computed directly in any straightforward Monte Carlo or

Molecular Dynamics simulation by simply recording the histogram of visited  $\zeta'$  values. Things become less evident when the probability distribution of the random variable is not only multimodal, but the regions in between the maxima are characterized by very low probabilities so that whatever simulation gets stuck in one of the highly probable regions and, physically speaking, we are in the presence of a metastability. When this is the case, a brute force sampling of the histogram is no longer possible and one has to find, by cunning, alternative ways to proceed. As we will see in the following, the concept of conditioned or constrained probability will take us out of the difficulty.

To see that, we consider the condition  $\zeta(\mathbf{r}) = \zeta'$  as the constraint

$$\sigma(\mathbf{r}) = \hat{\zeta}(\mathbf{r}) - \zeta' = 0, \tag{91}$$

and we compare the conditional probability

$$\mathcal{P}(\mathbf{r}|\zeta') = \frac{\mathcal{P}(\mathbf{r}, \zeta')}{\mathcal{P}(\zeta')} = \frac{e^{-\beta V} \delta(\hat{\zeta} - \zeta')}{\int d\mathbf{r} e^{-\beta V} \delta(\hat{\zeta} - \zeta')} \tag{92}$$

with Equation (80) for the marginal probability density in the presence of the constraint  $\sigma$ . We find

$$\mathcal{P}(\mathbf{r}|\zeta') = \frac{|\mathbf{Z}|^{-\frac{1}{2}} \mathcal{P}_M^{constr}(\mathbf{r})}{\int d\mathbf{r} |\mathbf{Z}|^{-\frac{1}{2}} \mathcal{P}_M^{constr}(\mathbf{r})} = \frac{|\mathbf{Z}|^{-\frac{1}{2}}}{\langle |\mathbf{Z}|^{-\frac{1}{2}} \rangle_{\sigma}^{constr}} \frac{\mathcal{P}_M^{constr}(\mathbf{r})}{\int d\mathbf{r} \mathcal{P}_M^{constr}(\mathbf{r})}, \tag{93}$$

i.e., a way to sample the conditional probability of  $\mathbf{r}$  given  $\zeta'$  by unbiasing a constrained probability density. Now, even regions of the configurational space associated with very low probabilities can be efficiently sampled and the metastability problem is taken out. In particular as a kind of corollary to Equation (93), we have for any configurational observable  $\hat{\mathcal{O}}(\mathbf{r})$ ,

$$\langle \hat{\mathcal{O}}(\mathbf{r}) \rangle_{\zeta'}^{cond} = \frac{\langle |\mathbf{Z}|^{-\frac{1}{2}} \hat{\mathcal{O}}(\mathbf{r}) \rangle_{\sigma}^{constr}}{\langle |\mathbf{Z}|^{-\frac{1}{2}} \rangle_{\sigma}^{constr}}, \tag{94}$$

with the rhs that, at variance with its left counterpart, can be efficiently sampled even for values of  $\zeta'$  corresponding to metastabilities.

The next problem arises when we need to take conditional averages for observables depending on the whole phase space, i.e.,  $\mathbf{r}$  and  $\mathbf{p}$ . This case, apparently not so common, is instead general if one considers conditional dynamic properties (time correlation functions) even of configurational properties. Indeed, as it is immediately evident,  $\hat{\mathcal{O}}(\mathbf{r}(t))$  is nothing else than a function of the initial condition  $(\mathbf{r}, \mathbf{p})$  parametrically dependent on the time  $t$ . Therefore, to be able to sample an unbiased conditional ensemble, with  $\sigma = 0$ , we need to have unbiased  $\mathcal{P}(\mathbf{r}, \mathbf{p}|\zeta')$ . We know that the momenta in the constrained ensemble, Equation (83), are irreversibly biased and thus unusable. However, we can unbias the configurations taken along a constrained trajectory and associate with them momenta sampled from an unbiased probability distribution. Knowing that, in the original ensemble, positions and momenta are independent and moreover the distribution of momenta  $\mathcal{P}_p(\mathbf{p})$  is just a product of Maxwellians, we can easily get such a sample

$$\mathcal{P}(\mathbf{r}, \mathbf{p}|\zeta') = \mathcal{P}_p(\mathbf{p}) \mathcal{P}_r(\mathbf{r}|\zeta') \tag{95}$$

from which directly a computable expression for a time correlation function at given  $\zeta'$

$$\langle \hat{\mathcal{O}}(\mathbf{r}(t)) \hat{\mathcal{O}}(\mathbf{r}) | \zeta' \rangle = \frac{\langle |\mathbf{Z}|^{-\frac{1}{2}} \hat{\mathcal{O}}(\mathbf{r}(t)) \hat{\mathcal{O}}(\mathbf{r}) | \zeta' \rangle_{\sigma}^{constr}}{\langle |\mathbf{Z}|^{-\frac{1}{2}} \rangle_{\sigma}^{constr}}, \tag{96}$$

where the time evolution now has to be intended to be fully unconstrained. The ensemble so constructed is the Blue Moon Ensemble and the problem of this particular metastability is now solved. In particular,

if we are interested in the calculation of an unconditioned time correlation function in a system where a brute force calculation (due to metastability) is not possible, we can compute it by thermodynamic integration using the predetermined marginal probability of  $\xi$ ,  $\mathcal{P}_\xi(\xi')$ . We get

$$\langle \hat{O}(\mathbf{r}(t)) \hat{O}(\mathbf{r}) \rangle = \int d\xi' \mathcal{P}_\xi(\xi') \langle \hat{O}(\mathbf{r}(t)) \hat{O}(\mathbf{r}) | \xi' \rangle. \tag{97}$$

To simplify the algebra and the notation, we have developed our argument only in the case of one scalar condition,  $\xi$ . The generalization to a vectorial condition is straightforward but cumbersome, and it can also be found explicitly derived in the literature [29]. The case in which the mechanical system contains constraints other than the ones representing physical conditions, typically molecular constraints, is formally more involved but conceptually identical. The interested reader can find all needed details in Ciccotti et al. [2].

### 6. Liouville Equation in the Presence of Constraints

The careful reader will have noticed at this point that we have properly solved the dynamics of a Hamiltonian system subjected to holonomic constraints and also formulated, in the Cartesian space, its statistical behavior at equilibrium. Instead, we have been unable to formulate in general the Statistical Mechanics of the system, including the evolution of the non equilibrium ensemble. The reason is that we miss the Liouville equation for this family of dynamical systems. We will see in the following that a generalized Liouville equation, always in the Cartesian reference description, can be derived at the price of abandoning the formulation of the dynamical evolution of the system by Lagrange multipliers and deriving, instead, the statistical behavior of a many-body, non-Hamiltonian system still satisfying the (assumed) conditions needed to justify a statistical treatment (e.g., chaotic behavior of the constituents, etc.). In these conditions, we will be able

- (i) to get a correct generalized Liouville equation;
- (ii) to find the results already obtained for the equilibrium ensemble;
- (iii) to generalize to these systems the theory of the response to external perturbations (in particular Kubo linear response theory [30]) well known and of widespread use for Hamiltonian systems, not only theoretically but also in molecular dynamics simulations [19,31–33].

A general, non-Hamiltonian, dynamical autonomous system is defined, in the set of variables  $\{\mathbf{x}\} (\equiv \{\mathbf{r}, \mathbf{p} = m\dot{\mathbf{r}}\})$ , by

$$\dot{\mathbf{x}} = \mathbf{G}(\mathbf{x}) \tag{98}$$

with the single component  $G_i(\mathbf{x})$  not derivable as  $\pm \frac{\partial \mathcal{H}}{\partial x_i}$ . The first and most important difference with a Hamiltonian system, especially in view of the derivation of the statistical properties of such a system, is that the phase space volume can be no more an invariant of the motion. If that happens, the standard approach of Statistical Mechanics is doomed to fail. However, it is easy to see that an invariant measure for the systems given in Equation (98) is easily found. Indeed, now

$$d\mathbf{x}_t = J(\mathbf{x}_t, \mathbf{x}_0) d\mathbf{x}_0, \tag{99}$$

where  $\mathbf{x}_t = \mathbf{x}_t(t; \mathbf{x}_0)$  is the solution of Equation (98),  $\mathbf{x}_0$  the initial condition and  $J_t = J(\mathbf{x}_t, \mathbf{x}_0)$  the determinant of the Jacobian matrix of the time-generated change of variables. In Appendix B, it is shown that

$$\frac{d}{dt} J(\mathbf{x}_t, \mathbf{x}_0) = \left( \nabla_{\mathbf{x}_t}^T \dot{\mathbf{x}}_t \right) J(\mathbf{x}_t, \mathbf{x}_0) = \kappa(\mathbf{x}_t) J(\mathbf{x}_t, \mathbf{x}_0), \tag{100}$$

where  $\kappa$ , the divergence of the flow in phase space  $\kappa(\mathbf{x}_t) = \nabla_{\mathbf{x}_t}^T \dot{\mathbf{x}}_t$ , is known as the phase space compressibility of the dynamical system, and we have introduced the shorthand notation  $\nabla_{\mathbf{x}_t}$  for the

gradient  $\nabla_x = \frac{\partial}{\partial x}$  with respect to coordinates  $x_t$  at time  $t$ . The solution of Equation (100) with the initial condition  $J(x_{t=0}, x_0) = 1$  is

$$J(x_t, x_0) = \exp \left\{ \int_0^t d\tau \kappa(x_\tau) d\tau \right\} = \exp \{ w(x_t, t) - w(x_0, 0) \}, \tag{101}$$

where  $w(x_t, t)$  is the primitive function associated with the indefinite time integral of  $\kappa(x_t)$ , which exists with certainty given that  $\kappa(x_t) = \frac{d \ln J_t}{dt}$ . Substituting this results in Equation (99), we find

$$\exp \{ -w(x_t, t) \} dx_t = \exp \{ -w(x_0, 0) \} dx_0 \tag{102}$$

i.e., the conservation in time of the measure  $e^{-w(x_t, t)} dx_t$ . The factor  $\exp \{ -w(x_t, t) \}$ , let us call it  $\gamma(x_t, t)$ , is the metric factor associated with the coordinate transformation  $x_0 \rightarrow x_t$ . It tells us that the statistical space of the variables  $x$  is no more Euclidean but has a non trivial metric structure. Remembering that the statistical ensemble is described by a probability density  $\mathcal{P}(x, t)$ , we have for the normalization condition,

$$\int dx \gamma(x) \mathcal{P}(x, t) = 1, \quad \forall t, \tag{103}$$

from which we get  $\frac{d}{dt} (\int dx \gamma \mathcal{P}) = 0$  and, therefore, the continuity equation

$$\frac{\partial}{\partial t} (\gamma(x) \mathcal{P}(x, t)) + \nabla_x^T (\dot{x} \gamma(x) \mathcal{P}(x, t)) = 0, \tag{104}$$

which represents the new, valid form for the Liouville equation for our more general dynamical systems. The solutions of Equation (104) will give us the evolution in the time of the ensemble associated with our non-Hamiltonian systems in non-equilibrium conditions while their stationary, asymptotic solutions can represent the equilibrium ensemble. We will now proceed in two steps, in order to derive the consequences of this more general approach to Statistical Mechanics. First, we will discuss how to obtain, in these conditions, not just a stationary solution but the correct equilibrium ensemble corresponding to the microcanonical ensemble of the standard case and show, just for illustration, how by this procedure we can re-derive the equilibrium ensemble for systems subjected to holonomic constraints. Then, second, using the general form of the solution of Equation (104), we will show the validity of the response approach developed by Onsager and Kubo (at least) for the Hamiltonian case.

### 6.1. Generalized Distribution Function

Assuming that the system (98) possesses  $n_c$  conserved quantities  $\hat{C}_k(x)$ ,  $k = 1, \dots, n_c$ ,

$$\frac{d\hat{C}_k(x)}{dt} = 0, \quad k = 1, \dots, n_c, \tag{105}$$

the space sampled by its trajectories will be the subspace intersection of the hypersurfaces  $\hat{C}_k(x) = c_k$ , where the values  $c_k$  are determined by the initial conditions. The “microcanonical” distribution function generated in these conditions is

$$\mathcal{P}^{(micro)}(x) \propto \prod_{k=1}^{n_c} \delta(\hat{C}_k(x) - c_k). \tag{106}$$

The solution (106) satisfies Equation (104) since its total time derivative is evidently zero and, moreover, is microcanonical since all accessible configurations are equiprobable. Other solutions exist, for example products of delta functions for subsets of the full set of conservation laws, but they do not

correspond to physical ensembles since they will represent hypersurfaces containing states that will never be visited. In other words, physical ensembles cannot be obtained by using only the solutions of the Liouville equation (104). To satisfy the stationary Liouville equation is a necessary but not sufficient condition. From the previous observations, it is possible to derive the rules to be followed to construct the proper equilibrium ensemble and the correct invariant measure.

The ensemble:

1. Construct the distribution function by Equation (106) using all the independent conservation laws implicit in the equations of motion;
2. Eliminate from the statistical space all variables that result uncoupled to the bulk of the system or driven by it. By driven, we mean variables
  - (i) whose evolution follows that of the other variables without influencing those ones and
  - (ii) that do not appear in the phase space expression of any of the  $n_c$  conserved quantities  $\hat{C}_k$ .

A (not so) typical example could be that of particles of zero mass interacting with the system only via the holonomic constraints defining their own values (see Appendix C).

The measure:

3. Once the essential, reduced, set of variables, let us call them  $\mathbf{x}'$ , has been selected, calculate the phase space compressibility  $\kappa(\mathbf{x}') = \nabla_{\mathbf{x}'}^T \dot{\mathbf{x}}'$  of the reduced dynamical system

$$\dot{\mathbf{x}}' = \tilde{\mathbf{G}}(\mathbf{x}') \tag{107}$$

and use  $\kappa(\mathbf{x}')$  to determine  $\gamma(\mathbf{x}')$ .

The results are

$$\mathcal{P}^{(micro)}(\mathbf{x}') = \frac{\prod_{k=1}^{n_c} \delta(\hat{C}_k(\mathbf{x}') - c_k)}{\int d\mathbf{x}' \gamma(\mathbf{x}') \prod_{k=1}^{n_c} \delta(\hat{C}_k(\mathbf{x}') - c_k)}, \tag{108}$$

where, via the normalization factor, is implicitly defined the new partition function

$$\Omega(c_1, \dots, c_{n_c}) = \int d\mathbf{x}' \gamma(\mathbf{x}') \prod_{k=1}^{n_c} \delta(\hat{C}_k(\mathbf{x}') - c_k). \tag{109}$$

We now turn, for illustrative purposes, to apply the formalism just developed to an originally Hamiltonian dynamical system subjected to holonomic constraints. As we have seen before, the non-Hamiltonian equations of motion are obtained inserting directly Equation (10) into Equation (14). The result is

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}, \tag{110}$$

$$\begin{aligned} \dot{\mathbf{p}}_i = & -\frac{\partial \mathcal{V}}{\partial \mathbf{r}_i} - \sum_{\alpha=1}^f \left\{ \sum_{\theta=1}^f \left[ -\frac{1}{m_i} \frac{\partial \mathcal{V}}{\partial \mathbf{r}_i} \cdot \left( \frac{\partial}{\partial \mathbf{r}_i} \right) \sigma_{\theta} \right. \right. \\ & \left. \left. + \sum_{k=1}^N \sum_{j=1}^N \left( \frac{\mathbf{p}_k}{m_k} \cdot \frac{\partial}{\partial \mathbf{r}_k} \right) \left( \frac{\mathbf{p}_j}{m_j} \cdot \frac{\partial}{\partial \mathbf{r}_j} \right) \sigma_{\theta} \right] (\mathbf{Z}^{-1})_{\theta, \alpha} \right\} \frac{\partial \sigma_{\alpha}(\mathbf{r})}{\partial \mathbf{r}_i}, \quad i = 1, \dots, N. \end{aligned} \tag{111}$$

Now,  $\sigma_\alpha = \dot{\sigma}_\alpha = 0, \alpha = 1, \dots, f$  are  $2f$  conservation laws to be added to the Hamiltonian. The compressibility factor  $\kappa$  is (easily) computed as

$$\kappa = \nabla_x^T \dot{x} = \sum_{i=1}^N \left( \frac{\partial}{\partial p_i} \cdot \dot{p}_i \right) = - \sum_{\alpha=1}^f \sum_{i=1}^N \left( \frac{\partial \lambda_\alpha}{\partial p_i} \right) \cdot \frac{\partial \sigma_\alpha}{\partial r_i} \tag{112}$$

$$= -2 \sum_{\alpha=1}^f \sum_{\theta=1}^f \sum_{i=1}^N \left[ \left( \frac{1}{m_i} \frac{\partial}{\partial r_i} \right) \cdot \left( \sum_{j=1}^N \frac{p_j}{m_j} \cdot \frac{\partial \sigma_\theta}{\partial r_j} \right) \right] (\mathbf{Z}^{-1})_{\theta,\alpha} \frac{\partial \sigma_\alpha}{\partial r_i} \tag{113}$$

$$= - \sum_{\alpha=1}^f \sum_{\theta=1}^f \left[ 2 \sum_{i=1}^N \frac{1}{m_i} \frac{\partial \sigma_\alpha}{\partial r_i} \cdot \frac{\partial \dot{\sigma}_\theta}{\partial r_i} \right] (\mathbf{Z}^{-1})_{\theta,\alpha} = - \sum_{\alpha=1}^f \sum_{\theta=1}^f \dot{Z}_{\alpha,\theta} (\mathbf{Z}^{-1})_{\theta,\alpha} = - \frac{d}{dt} \ln |\mathbf{Z}|, \tag{114}$$

giving

$$\gamma = e^{-w} = |\mathbf{Z}|, \tag{115}$$

from which we recover the ensemble already derived, Equation (73).

### 6.2. Response Theory

We address now the central question of dynamical non-equilibrium Statistical Mechanics for systems subjected to holonomic constraints: how to get statistical averages when the evolution of the system is no more stationary be it due to time-dependent perturbations or to the study of relaxation processes [8]. These problems are already solved in the Hamiltonian case (even with non-Hamiltonian perturbations but conserving the phase space volume [19,31–35]); here, we extend that solution to our present case. Let us start from the simpler case of the study of relaxation. Here, we have the system prepared in a non-equilibrium condition and we intend to study the macroscopic relaxation of an observable:

$$\langle \hat{O} \rangle_t = \int d\omega(\mathbf{x}) \hat{O}(\mathbf{x}) \mathcal{P}(\mathbf{x}, t), \tag{116}$$

where  $d\omega(\mathbf{x}) = \gamma(\mathbf{x})d\mathbf{x}$  is the invariant measure already derived and  $\mathcal{P}(\mathbf{x}, t)$  is the ensemble at time  $t$  obtained by evolving with the generalized Liouville Equation (104) the initial non-stationary ensemble  $\mathcal{P}(\mathbf{x}, 0) \equiv P(\mathbf{x})$ . Let us define the Liouville operator

$$i\hat{\mathcal{L}}_0 = \dot{x}^T \nabla_x = \boldsymbol{\mu}^{-1} \mathbf{p}^T \nabla_r - (\nabla_x \mathcal{V})^T \nabla_p - \sum_{\alpha=1}^f \lambda_\alpha(\mathbf{x}) (\nabla_r \sigma_\alpha)^T \nabla_p. \tag{117}$$

It follows immediately that

$$\frac{d\hat{O}(\mathbf{x}(t))}{dt} = i\hat{\mathcal{L}}_0 \hat{O}(\mathbf{x}(t)) \tag{118}$$

with formal solution

$$\hat{O}(\mathbf{x}(t)) = e^{i\hat{\mathcal{L}}_0 t} \hat{O}(\mathbf{x}). \tag{119}$$

As for the evolution of the ensemble, we can start from the Liouville equation (104)

$$\frac{\partial \gamma \mathcal{P}}{\partial t} = - \nabla_x^T (\dot{x} \gamma \mathcal{P}) = - (i\hat{\mathcal{L}}_0 + \kappa) (\gamma \mathcal{P}), \tag{120}$$

from which we can (easily) find, by using the fact that  $|\mathbf{Z}| = |\mathbf{Z}(\mathbf{r})|$  and the identity  $|\mathbf{Z}| = e^{\text{Tr} \ln \mathbf{Z}}$ , that

$$i\hat{\mathcal{L}}_0\gamma = \sum_{i=1}^N \frac{\mathbf{p}_i}{m_i} \cdot \frac{\partial |\mathbf{Z}|}{\partial \mathbf{r}_i} = \sum_{\alpha=1}^f \sum_{\theta=1}^f \left[ \sum_{i=1}^N \frac{\mathbf{p}_i}{m_i} \cdot \frac{\partial Z_{\alpha\theta}}{\partial \mathbf{r}_i} \right] (\mathbf{Z}^{-1})_{\theta\alpha} |\mathbf{Z}| \tag{121}$$

$$= \left[ \sum_{\alpha=1}^f \sum_{\theta=1}^f \dot{Z}_{\alpha\theta} (\mathbf{Z}^{-1})_{\theta\alpha} \right] |\mathbf{Z}| = -\kappa \gamma, \tag{122}$$

i.e.,

$$\frac{\partial \mathcal{P}}{\partial t} = (-i\hat{\mathcal{L}}_0) \mathcal{P}, \tag{123}$$

or else, once again,

$$\frac{d\mathcal{P}}{dt} = \frac{\partial \mathcal{P}}{\partial t} + i\hat{\mathcal{L}}_0\mathcal{P} = 0. \tag{124}$$

Moreover, as in the standard case

$$\mathcal{P}(\mathbf{x}, t) = e^{-i\hat{\mathcal{L}}_0 t} \mathcal{P}(\mathbf{x}). \tag{125}$$

By remembering that we are working with an invariant measure, we find, again,

$$\langle \hat{\mathcal{O}} \rangle_t = \int d\omega \hat{\mathcal{O}}(\mathbf{x}) \mathcal{P}(\mathbf{x}, t) = \int d\omega \left[ e^{i\hat{\mathcal{L}}_0 t} \hat{\mathcal{O}}(\mathbf{x}) \right] P(\mathbf{x}) = \int d\omega \hat{\mathcal{O}}(\mathbf{x}(t)) P(\mathbf{x}), \tag{126}$$

a relation easy to implement in molecular dynamics simulation of a relaxation process if we can prepare a sample of the non-equilibrium initial ensemble  $P(\mathbf{x})$ .

Let us now move to the case in which we are interested to compute the response of the system to an external time-independent field. The equations of motion become

$$\dot{\mathbf{x}} = \mathbf{G}(\mathbf{x}) + \mathbf{D}(\mathbf{x})\mathcal{F}(t), \tag{127}$$

with  $\mathbf{D}(\mathbf{x})$  derivable ( $\mathbf{D}(\mathbf{r}) = \frac{\partial H_p}{\partial \mathbf{p}}$ ;  $\mathbf{D}(\mathbf{p}) = -\frac{\partial H_p}{\partial \mathbf{r}}$ ) or not from a Hamiltonian perturbation term ( $H_p$ ). In any event, to simplify the formalism (and on the basis of what is usually done in transport studies [36]), let us assume that  $\nabla_{\mathbf{x}}^T \mathbf{D}(\mathbf{x}) = 0$ , i.e., that the perturbation satisfies the incompressibility condition. This condition guarantees that, even in the presence of the perturbation, the non-zero compressibility arises only from the constraints and it is given by Equation (114). The Liouville operator is now time-dependent

$$i\hat{\mathcal{L}}(t) = \mathbf{p}^T \nabla_{\mathbf{r}} - \left( \nabla_{\mathbf{r}}^T \mathcal{V} + \sum_{\alpha} \lambda_{\alpha} \nabla_{\mathbf{r}}^T \sigma_{\alpha} \right) \nabla_{\mathbf{p}} + \mathcal{F}(t) \mathbf{D}^T \nabla_{\mathbf{x}} \tag{128}$$

and

$$\hat{\mathcal{O}}(\mathbf{x}(t)) = \mathcal{T} \exp \left\{ i \int_0^t d\tau \hat{\mathcal{L}}(\tau) \right\} \hat{\mathcal{O}}(\mathbf{x}), \tag{129}$$

where  $\mathcal{T}$  is the time-ordering operator. However, in spite of this more daring complexity, again the probability density evolves with the operator

$$S^{\dagger}(t, 0) = \mathcal{T} \exp \left\{ -i \int_0^t d\tau \hat{\mathcal{L}}(\tau) \right\} \tag{130}$$

so that, again,

$$\langle \hat{O} \rangle_t = \int d\omega \hat{O}(\mathbf{x}) S^\dagger(t, 0) P(\mathbf{x}) \tag{131}$$

$$= \int d\omega \hat{O}(\mathbf{x}(t)) P(\mathbf{x}) \tag{132}$$

is the initial condition an equilibrium distribution or a general one, in any event a relation easy to implement in molecular dynamics simulations. Equation (132) is valid in general both in the linear response regime and beyond it. In the case of small perturbations, it is possible to show, after some algebra, that, in the presence of constraints, the classical linear response result of Green [37] and Kubo [30] is recovered and holds without any alteration from the unconstrained case [9].

### 7. Conclusions

The dynamics and Statistical Mechanics of a many-body system subjected to holonomic constraints have been discussed both following, as for the equilibrium case, the classical historical Lagrange (Hamilton) approach, using Lagrangian multipliers, and, more generally, from the newer perspective, encompassing also non-equilibrium, of non-Hamiltonian flows in phase space. One section has been dedicated to review in depth the most relevant numerical implementations, while, for the sake of readability, the reader has been addressed to the relevant literature for the technically most involved cases. A quite peculiar application, the zero-mass particle case, has been discussed to show how constraints can be creatively used to extend the description of the system opening a different, efficient, way of incorporating, for example, new features in force field models. Let us, finally, remark that developing the statistical theory of dynamical systems subjected to holonomic constraints permits to cover both equilibrium and non-equilibrium simulations of molecular systems but also to explore the domain of rare events, including the computing of complex free energy landscapes, the probing of the dynamics of rare events and even performing non-equilibrium hydrodynamical simulations by properly sampling initial conditions assigning the proper weight to the ensemble of non-equilibrium trajectories that gives the correct (linear and nonlinear) response.

**Acknowledgments:** It gives us great pleasure to acknowledge the importance of the lifelong collaboration on these themes with Jean-Paul Ryckaert, Ray Kapral, Michiel Sprik, Glenn Martyna, Mark Tuckerman, Eric Vanden-Eijnden and Carsten Hartmann. This work was partially supported by grant MIUR PRIN-2012NNRCAF\_004.

**Author Contributions:** The authors contributed equally to this work.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

To evaluate the thermodynamic force in Equation (90), we need to separately address the contributions from the kinetic and the potential terms in the Hamiltonian  $\mathcal{H}^*(\mathbf{u}, \mathbf{p}_u)$ ,

$$\frac{d\mathcal{W}_\xi(\xi')}{d\xi'} = \frac{\int d\mathbf{u} d\mathbf{p}_u \left[ \left( \frac{\partial \mathcal{K}^*}{\partial \xi} \right) + \left( \frac{\partial \mathcal{V}^*}{\partial \xi} \right) \right] e^{-\beta \mathcal{H}^*} \delta(\xi - \xi')}{\int d\mathbf{u} d\mathbf{p}_u e^{-\beta \mathcal{H}^*} \delta(\xi - \xi')}. \tag{A1}$$

The potential term does not depend on the momenta  $\mathbf{p}_u$  and its contribution is immediately addressed. The kinetic term  $\mathcal{K}^* = \frac{1}{2} \mathbf{p}_u^T \mathbf{M}^{-1} \mathbf{p}_u$  can be rearranged in a more significant form by performing the integral over the momenta,

$$\frac{\int d\mathbf{p}_u \left[ \frac{1}{2} \mathbf{p}_u^T \frac{\partial \mathbf{M}^{-1}}{\partial \xi} \mathbf{p}_u \right] \exp\left\{-\frac{\beta}{2} \mathbf{p}_u^T \mathbf{M}^{-1} \mathbf{p}_u\right\}}{\int d\mathbf{p}_u \exp\left\{-\frac{\beta}{2} \mathbf{p}_u^T \mathbf{M}^{-1} \mathbf{p}_u\right\}} \tag{A2}$$

$$= -k_B T \frac{\frac{\partial}{\partial \xi} \int d\mathbf{p}_u \exp\left\{-\frac{\beta}{2} \mathbf{p}_u^T \mathbf{M}^{-1} \mathbf{p}_u\right\}}{\int d\mathbf{p}_u \exp\left\{-\frac{\beta}{2} \mathbf{p}_u^T \mathbf{M}^{-1} \mathbf{p}_u\right\}} \tag{A3}$$

$$= -k_B T \frac{\frac{\partial}{\partial \xi} \left(\frac{2\pi}{\beta}\right)^{3N/2} |\mathbf{M}|^{1/2}}{\left(\frac{2\pi}{\beta}\right)^{3N/2} |\mathbf{M}|^{1/2}} \tag{A4}$$

$$= -\frac{k_B T}{2} \left(\frac{\partial \ln |\mathbf{M}|}{\partial \xi}\right) = -k_B T \left(\frac{\partial \ln J}{\partial \xi}\right), \tag{A5}$$

where we have used Equation (77) and Equation (21).

### Appendix B

We enclose for self-consistency of our text a straightforward demonstration of Equation (100). If the coordinate system is changed from coordinate  $\xi = x(0)$  to the coordinates  $x = x(t|\xi)$  from the time evolution transformation, the volume element in the  $n$ -dimensional space changes accordingly to the formula

$$dx_1 dx_2 \dots dx_n = dx = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(\xi_1, \xi_2, \dots, \xi_n)} d\xi_1 d\xi_2 \dots d\xi_n = J d\xi, \tag{A6}$$

where  $J = |J|$  is the Jacobian, i.e., the determinant of the Jacobian matrix

$$(J)_{i,k} = \frac{\partial x_i}{\partial \xi_k}. \tag{A7}$$

A simple linear algebra result, based on the invariance of determinants and traces with respect to unitary transformations, such as diagonalization, states that, for a given square matrix  $A$ , with  $\tilde{A} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$  its diagonal form, the determinant  $|A| = |\tilde{A}| = \prod_i (\tilde{A})_{i,i}$  can be expressed using  $\text{Tr} \ln A$ , the trace of the logarithm of the matrix itself:

$$e^{\text{Tr} \ln A} = e^{\text{Tr} \ln \tilde{A}} = e^{\sum_i \ln(\tilde{A})_{i,i}} = \prod_i (\tilde{A})_{i,i} = |A|. \tag{A8}$$

Applying Equation (A8) to the matrix  $J$ , one obtains for the Jacobian the formula

$$J(x, \xi) = |J| = e^{\text{Tr} \ln J}. \tag{A9}$$

Equation (A9) can be derived with respect to time  $t$  to obtain for  $J$  the equation of motion

$$\frac{dJ}{dt} = \frac{d e^{\text{Tr} \ln J}}{dt} = \text{Tr} \left( \frac{dJ}{dt} J^{-1} \right) e^{\text{Tr} \ln J} = \left( \sum_{i=1}^n \sum_{k=1}^n \left( \frac{dJ}{dt} \right)_{i,k} (J^{-1})_{k,i} \right) J. \tag{A10}$$

The derivatives of the elements of the Jacobian matrix can be expressed in terms of the velocity field  $\dot{x} = \frac{dx}{dt} = G(x)$ , where we have used Equation (98) to remind readers that the “velocities”  $\dot{x}_i$ ,  $i = 1, \dots, n$  can be expressed as functions of the coordinates  $x(\xi)$  and, therefore, by exchanging the order of derivation,

$$\left( \frac{dJ}{dt} \right)_{i,k} = \frac{d}{dt} \frac{\partial x_i}{\partial \xi_k} = \frac{\partial \dot{x}_i}{\partial \xi_k}. \tag{A11}$$

Substituting Equation (A11) in Equation (A10), and reminding readers that  $(J^{-1})_{k,i} = \frac{\partial \zeta_k}{\partial x_i}$ , we finally obtain

$$\frac{dJ}{dt} = \left( \sum_{i=1}^n \left[ \sum_{k=1}^n \frac{\partial \dot{x}_i}{\partial \zeta_k} \frac{\partial \zeta_k}{\partial x_i} \right] \right) J = \left( \sum_{i=1}^n \frac{\partial \dot{x}_i}{\partial x_i} \right) J = \kappa J, \tag{A12}$$

i.e., Equation (100).

### Appendix C

The axis of a diatomic molecule cannot provide a reference frame (a comoving frame) attached to the molecule. However, adding to the molecule a third point of mass zero not collinear with the physical molecule, we can get the comoving frame we were looking for. We show below that this extra variable is driven and doesn't alter the dynamics and Statistical Mechanics of our system [38]. In order to create a rigid triatomic molecule, i.e., a rigid triangle, one needs to specify three so-called "bond" constraints for the square distances between each pair of atoms

$$\sigma_{ij} = \frac{1}{2} \left[ (\mathbf{r}_j - \mathbf{r}_i)^2 - d_{ij}^2 \right], \quad ij = 12, 23, 31, \tag{A13}$$

where  $\mathbf{r}_i$  is the three-dimensional atomic coordinate of atom  $i$  with  $i = 1, 2, 3$  and  $d_{ij}$  is the (rigid) distance between atoms  $i$  and  $j$ . Bond constraints are easy to deal with since, as  $\frac{\partial \sigma_{ij}}{\partial \mathbf{r}_i} = -(\mathbf{r}_j - \mathbf{r}_i)$ , the constraint force is parallel to the bond, and the equations of motion are

$$m_1 \ddot{\mathbf{r}}_1 = \mathbf{F}_1 + \lambda_{12}(\mathbf{r}_2 - \mathbf{r}_1) + \lambda_{31}(\mathbf{r}_3 - \mathbf{r}_1), \tag{A14}$$

$$m_2 \ddot{\mathbf{r}}_2 = \mathbf{F}_2 + \lambda_{23}(\mathbf{r}_3 - \mathbf{r}_2) + \lambda_{12}(\mathbf{r}_1 - \mathbf{r}_2), \tag{A15}$$

$$m_3 \ddot{\mathbf{r}}_3 = \mathbf{F}_3 + \lambda_{31}(\mathbf{r}_1 - \mathbf{r}_3) + \lambda_{23}(\mathbf{r}_2 - \mathbf{r}_3), \tag{A16}$$

where  $\mathbf{F}_i$  is the total force acting on atom  $i$ . Assuming  $i = 3$  is the index of the virtual particle and  $\mathbf{F}_3 = 0$ , one immediately has that

$$m_3 \ddot{\mathbf{r}}_3 = \lambda_{31}(\mathbf{r}_1 - \mathbf{r}_3) + \lambda_{23}(\mathbf{r}_2 - \mathbf{r}_3) = 0, \tag{A17}$$

implying  $\lambda_{31} = 0$  and  $\lambda_{23} = 0$ , as the two non-zero bond vectors  $\mathbf{r}_1 - \mathbf{r}_3$  and  $\mathbf{r}_2 - \mathbf{r}_3$  are, by definition, not collinear. By taking the limit for  $m_3 \rightarrow 0$  of Equation (A16) after dividing it by  $m_3$  one has that, although the total force acting of  $\mathbf{r}_3$  is zero, the acceleration

$$\ddot{\mathbf{r}}_3 = \zeta_1(\mathbf{r}_1 - \mathbf{r}_3) + \zeta_2(\mathbf{r}_2 - \mathbf{r}_3), \quad \left( \zeta_1 = \lim_{m_3 \rightarrow 0} \frac{\lambda_{31}}{m_3}, \quad \zeta_2 = \lim_{m_3 \rightarrow 0} \frac{\lambda_{23}}{m_3} \right) \tag{A18}$$

does not need to vanish, and the equations of motion can be rewritten as

$$\begin{cases} m_1 \ddot{\mathbf{r}}_1 = \mathbf{F}_1 + \lambda_{12}(\mathbf{r}_2 - \mathbf{r}_1), \\ m_2 \ddot{\mathbf{r}}_2 = \mathbf{F}_2 + \lambda_{12}(\mathbf{r}_1 - \mathbf{r}_2), \\ \ddot{\mathbf{r}}_3 = \zeta_1(\mathbf{r}_1 - \mathbf{r}_3) + \zeta_2(\mathbf{r}_2 - \mathbf{r}_3). \end{cases} \tag{A19}$$

One can notice that, as expected, the motion of the two "real" atoms is not affected by the addition of the virtual one and the dynamics of the third virtual atom simply follows the motion of the first two, "driven" by the constraints. Noting, moreover, that the new variable doesn't enter in any of the conservation laws of the system, we can conclude safely that also the statistical behavior of the system is not altered by the presence of the extra particle. Equations (A19) can be integrated numerically using SHAKE with the Verlet algorithm described in Section 3.

Another, possibly more interesting, case can arise with a force-field model containing extra-centers of force, whose positions do not coincide with the atomic positions but follow adiabatically the motion of the atoms, taking positions that satisfy the condition of zero force on them. The dynamics of these extra zero-mass points are again inherently driven by that of the material points and so does not intervene in the statistical behavior of the material system.

## References

1. Carter, E.; Ciccotti, G.; Hynes, J.T.; Kapral, R. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.* **1989**, *156*, 472–477.
2. Ciccotti, G.; Kapral, R.; Vanden-Eijnden, E. Blue Moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem* **2005**, *6*, 1809–1814.
3. Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
4. Car, R.; Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
5. Ciccotti, G.; Ferrario, M. Constrained and nonequilibrium molecular dynamics. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; World Scientific: Singapore, 1998; pp. 157–177.
6. Ryckaert, J.P.; Ciccotti, G. Introduction of Andersen's demon in the molecular dynamics of systems with constraints. *J. Chem. Phys.* **1983**, *78*, 7368–7374, doi:10.1063/1.444728.
7. Tuckerman, M.E.; Liu, Y.; Ciccotti, G.; Martyna, G.J. Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems. *J. Chem. Phys.* **2001**, *115*, 1678–1702, doi:10.1063/1.1378321.
8. Ciccotti, G.; Kapral, R.; Sergi, A. Non-equilibrium molecular dynamics. In *Handbook of Materials Modeling*; Yip, S., Ed.; Springer: Berlin, Germany, 2005; pp. 745–761.
9. Hartmann, C.; Schütte, C.; Ciccotti, G. Communications: On the linear response of mechanical systems with constraints. *J. Chem. Phys.* **2010**, *132*, 111103, doi:10.1063/1.3354126.
10. Goldstein, H.; Poole, C.P.; Safko, J.L. *Classical Mechanics*, 3rd Edition; Addison-Wesley: Boston, MA, USA, 2000.
11. Ciccotti, G.; Ryckaert, J. Molecular dynamics simulation of rigid molecules. *Comput. Phys. Rep.* **1986**, *4*, 346–392.
12. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.
13. Weinbach, Y.; Elber, R. Revisiting and parallelizing SHAKE. *J. Comput. Phys.* **2005**, *209*, 193–206.
14. Ciccotti, G.; Ferrario, M.; Hynes, J.T.; Kapral, R. Molecular dynamics simulation of ion association reactions in a polar solvent. *J. Chim. Phys.* **1988**, *85*, 925–929.
15. Sprik, M.; Ciccotti, G. Free energy from constrained molecular dynamics. *J. Chem. Phys.* **1998**, *109*, 7737–7744, doi:10.1063/1.477419.
16. Orlandini, S.; Meloni, S.; Ciccotti, G. Hydrodynamics from Statistical Mechanics: Combined dynamical-NEMD and conditional sampling to relax an interface between two immiscible liquids. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13177–13181.
17. Cottone, G.; Lattanzi, G.; Ciccotti, G.; Elber, R. Multiphoton absorption of myoglobin–nitric oxide complex: Relaxation by D-NEMD of a stationary state. *J. Phys. Chem. B* **2012**, *116*, 3397–3410, doi:10.1021/jp212148x, PMID: 22356468.
18. Pourali, M.; Meloni, S.; Magaletti, F.; Maghari, A.; Casciola, C.M.; Ciccotti, G. Relaxation of a steep density gradient in a simple fluid: Comparison between atomistic and continuum modeling. *J. Chem. Phys.* **2014**, *141*, 154107, doi:10.1063/1.4897977.
19. Ciccotti, G.; Ferrario, M. Dynamical non-equilibrium molecular dynamics. *Entropy* **2014**, *16*, 233–257.
20. Ciccotti, G.; Bonella, S.; Ferrario, M.; Pierleoni, C. Probabilistic derivation of spatiotemporal correlation functions in the hydrodynamic limit. *J. Phys. Chem. B* **2016**, *120*, 1996–2000, doi:10.1021/acs.jpcc.5b10066, PMID: 26720838.
21. Fixman, M. Classical Statistical Mechanics of constraints: A theorem and application to polymers. *Proc. Nat. Acad. Sci. USA* **1974**, *71*, 3050–3053.
22. Hairer, E.; Lubich, C.; Wanner, G. Geometric numerical integration illustrated by the Störmer–Verlet method. *Acta Numer.* **2003**, *12*, 399–450.

23. Hairer, E.; Wanner, G.; Lubich, C. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*; Springer: Berlin, Germany, 2006.
24. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
25. Krätler, V.; van Gunsteren, W.F.; Hünenberger, P.H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, *22*, 501–508.
26. Gonnet, P. P-SHAKE: A quadratically convergent SHAKE in  $O(n^2)$ . *J. Comput. Phys.* **2007**, *220*, 740–750.
27. Gonnet, P.; Walther, J.H.; Koumoutsakos, P.  $\theta$ -SHAKE: An extension to SHAKE for the explicit treatment of angular constraints. *Comput. Phys. Commun.* **2009**, *180*, 360–364.
28. Leimkuhler, B.; Reich, S. Symplectic integration of constrained Hamiltonian systems. *Math. Comput.* **1994**, *63*, 589–605.
29. Sergi, A.; Ciccotti, G.; Falconi, M.; Desideri, A.; Ferrario, M. Effective binding force calculation in a dimeric protein by molecular dynamics simulation. *J. Chem. Phys.* **2002**, *116*, 6329–6338, doi:10.1063/1.1457445.
30. Kubo, R. Statistical-Mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Japan* **1957**, *12*, 570–586, doi:10.1143/JPSJ.12.570.
31. Ciccotti, G.; Jacucci, G. Direct computation of dynamical response by molecular dynamics: The mobility of a charged Lennard-Jones particle. *Phys. Rev. Lett.* **1975**, *35*, 789–792.
32. Ciccotti, G.; Jacucci, G.; McDonald, I.R. “Thought-experiments” by molecular dynamics. *J. Stat. Phys.* **1979**, *21*, 1–22.
33. Ciccotti, G.; Ferrario, M. Non-equilibrium by molecular dynamics: A dynamical approach. *Mol. Simul.* **2016**, *42*, 1385–1400, doi:10.1080/08927022.2015.1121543.
34. Ferrario, M.; Bonella, S.; Ciccotti, G. On the establishment of thermal diffusion in binary Lennard-Jones liquids. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 1629–1642.
35. Bonella, S.; Ferrario, M.; Ciccotti, G. Thermal diffusion in binary mixtures: Transient behavior and transport coefficients from equilibrium and nonequilibrium molecular dynamics. *Langmuir* **2017**, *33*, 11281–11290, doi:10.1021/acs.langmuir.7b02565, PMID: 28915729.
36. Evans, D.J.; Morriss, G. *Statistical Mechanics of Nonequilibrium Liquids*; Cambridge University Press: Cambridge, UK, 2008.
37. Green, M.S. Markoff random processes and the Statistical Mechanics of time-dependent phenomena. *J. Chem. Phys.* **1952**, *20*, 1281–1295, doi:10.1063/1.1700722.
38. Ryckaert, J.P.; Bellemans, A.; Ciccotti, G. The rotation-translation coupling in diatomic molecules. *Mol. Phys.* **1981**, *44*, 979–996, doi:10.1080/00268978100102931.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).