Double density dynamics: realizing a joint distribution of a physical system and a parameter system

Ikuo Fukuda\textsuperscript{1} and Kei Moritsugu\textsuperscript{2}

\textsuperscript{1} Institute for Protein Research, Osaka University, Osaka 565-0871, Japan
\textsuperscript{2} Graduate School of Medical Life Science, Yokohama City University, Yokohama 230-0045, Japan

E-mail: ifukuda@protein.osaka-u.ac.jp

Received 31 May 2015
Accepted for publication 29 July 2015
Published 20 October 2015

Abstract
To perform a variety of types of molecular dynamics simulations, we created a deterministic method termed ‘double density dynamics’ (DDD), which realizes an arbitrary distribution for both physical variables and their associated parameters simultaneously. Specifically, we constructed an ordinary differential equation that has an invariant density relating to a joint distribution of the physical system and the parameter system. A generalized density function leads to a physical system that develops under nonequilibrium environment-describing superstatistics. The joint distribution density of the physical system and the parameter system appears as the Radon–Nikodym derivative of a distribution that is created by a scaled long-time average, generated from the flow of the differential equation under an ergodic assumption. The general mathematical framework is fully discussed to address the theoretical possibility of our method, and a numerical example representing a 1D harmonic oscillator is provided to validate the method being applied to the temperature parameters.

Keywords: ordinary differential equation, molecular dynamics, superstatistics, invariant measure, ergodicity, nonequilibrium simulation

(Some figures may appear in colour only in the online journal)
1. Introduction

Numerical simulations of physical systems are important tools for understanding the structures, phases, and mechanisms of the systems in a systematic manner. Molecular dynamics (MD) simulations are used for classical/quantum Hamiltonian/non-Hamiltonian systems to help understand the characteristics of the systems in terms of their microscopic descriptions [1].

The physical systems, or rather the equations of motion in these methods, include some parameters that may be key for performing the simulations. For example, intensive-quantity parameters such as temperature and pressure are important for understanding equilibrium and nonequilibrium states. There will be cases in which we are motivated to fluctuate these values to ascertain the effects of these parameters. A typical situation is to change or fluctuate these intensive quantity parameters, which is relevant for nonequilibrium simulations [2, 3]. It may also be useful to fluctuate the temperature to obtain a good sampling method [4, 5]. In other cases, there might be a parameter for which the best way to handle it is unclear. For example, the optimal values of the parameters are unknown a priori, such as the ‘mass’ parameters for the thermostat or barostat in extended system methods [6, 7]. In addition, the optimal parameter values in the sampling method are often system-dependent, and efforts are required for their identification. For example, the optimal value of the extensive parameter $q$ in the Tsallis-distribution sampling [8, 9] is unknown a priori.

In these situations, what should we do? Ad hoc methods for varying the parameter values do not provide important information, such as the probability distribution that the physical system obeys, and they thus preclude an easy interpretation to compare the results with experiments. The underlying motivation of the current study is to construct a route to solve this problem. A possible approach, which is the topic of this paper, is to control the dynamical variations of the parameter values by a certain protocol, based on a probabilistic description. Specifically, along the contexts introduced above, we consider performing the following three types of MD simulations. First, (I) temperature is fluctuated, obeying a certain probability distribution (for a nonequilibrium simulation or for good sampling). Second, taking account of a more general situation, where the target is not only the temperature parameter but also an arbitrary parameter, we consider that (II) any parameter is fluctuated, obeying a certain probability distribution (for varying an unknown parameter value). Third, considering the fact that such a fluctuating parameter yields a certain system other than the originally given physical system, a more general situation is that (III) any other system is fluctuated, obeying a certain probability distribution (for coupling two systems, e.g. all-atom model and the associated course-grained model [10]). Namely, our purpose is to produce a general deterministic equation of motion (EOM) in order to perform many kinds of MD simulations represented by these three categories (I)–(III), which are sorted in ascending order of generality.

How can we realize these methods? The key is to focus on rigorously connecting the probability densities and the dynamics that generate the densities. In this regard, we are fortunate to have a rich background to draw from. To create the Boltzmann–Gibbs (BG) (or canonical) distribution density described with the coordinate variables $x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n$ such as

$$\rho(x, \ldots) \propto \exp[-\beta U(x)]$$

via a deterministic EOM, the Gaussian isokinetic dynamics [11, 12] can be used. Here $U(x)$ is the potential energy of the physical system, and $\beta$ is the inverse temperature (Boltzmann’s constant is unity). Similarly, to create the BG distribution density described with both $x \in \mathbb{R}^n$...
and momentum variables \( p \equiv (p_1, \ldots, p_n) \in \mathbb{R}^n \) such as
\[
\rho(x, p, \ldots) \propto \exp[-\beta(U(x) + K(p))]
\] via a deterministic EOM, the Nosé–Hoover (NH) dynamics [13, 14] can be used. This is represented by
\[
\begin{align*}
\dot{x} &= M^{-1}p \in \mathbb{R}^n, \\
\dot{p} &= -\nabla U(x) - \gamma \zeta p \in \mathbb{R}^n, \\
\dot{\zeta} &= 2K(p) - n/\beta \in \mathbb{R},
\end{align*}
\] where \( \zeta \) is a variable to control the kinetic energy, \( K(p) \), by mimicking a function of a heat bath with a temperature of \( 1/\beta \), and \( \gamma \) is an associated parameter. The NH scheme has been actively utilized in revisiting thermostatting methods as detailed in a number of reviews [15, 16].

Based on this background, we will create an arbitrary distribution density of the form,
\[
\rho(x, p, Z, \ldots) \propto \rho_{\text{phys}}(x, p; Z) f(Z),
\] via a deterministic EOM, which is sought for. Here \( \rho_{\text{phys}}(x, p; Z) \) stands for a density value of a conditional probability for the physical system, such that the physical system variable \( (x, p) \) emerges under the condition that ‘parameter’ \( Z \) takes a certain value. \( f(Z) \) stands for a density value of the probability for the parameter itself. The left-hand side \( \rho(x, p, Z, \ldots) \) thus represents a joint distribution density defined in ‘physical-space × parameter-space.’ In the context of the three issues (I)–(III), this concept is specifically demonstrated, as follows. The situation that (I) ‘temperature \( 1/\beta \) is fluctuated obeying a certain probability distribution \( f' \)’ is described as a density,
\[
\rho(x, p, \beta, \ldots) \propto \rho_{\text{phys}}(x, p; \beta) f(\beta).
\] Considering the aforementioned conventional ideas, \( \rho_{\text{phys}}(x, p; \beta) \) corresponds to the rhs of equation (2). However, \( \beta \) in equation (5) is not a constant parameter but a dynamical variable obeying statistics characterized by the density, \( f \). Namely, we can imagine a physical view such that the target physical system interacts with a heat bath with a temperature that is instantaneously \( 1/\beta \), and this temperature itself fluctuates throughout the duration, according to the probability distribution density \( f \). The situation that (II) ‘any parameter \( q \) is fluctuated obeying a certain probability distribution \( f' \)’ is described as a density such as
\[
\rho(x, p, q, \ldots) \propto \rho_{\text{phys}}(x, p; q) f(q).
\] The situation that (III) ‘any other system described by \( X \) is fluctuated obeying a certain probability distribution \( f' \)’ is described as a density such as
\[
\rho(x, p, X, \ldots) \propto \rho_{\text{phys}}(x, p; X) f(X).
\] In this case, \( X \) may no longer be a ‘parameter,’ but it can be viewed as a generalized parameter concerned with the original physical system. Below, for simplicity, instead of using \( Z \) in the most general expression (4), we just use \( \beta \) as in the case of (I), and express the density as
\[
\rho_{\text{phys}}(x, p; \beta) f(\beta).
\] This is also because we shall mainly emphasize the temperature parameter in this paper, since temperature fluctuations are one of the most interesting and important issues in a variety of research areas. Namely, we realize a joint distribution density of the physical system and the parameter system, represented as the form (8). We derive an ordinary differential equation (ODE) that produces the density, equation (8), with dynamical variables including \( x, p, \) and
β, by providing an invariant density (a density of an invariant measure) that is related to equation (8).

The distribution of the physical system regarding equation (8) is represented by the marginal distribution density,

$$\rho_R(x, p) \equiv \int d\beta \rho_{\text{phys}}(x, p, \beta) f(\beta). \quad (9)$$

When we focus on situation (I) in which β is the (inverse) temperature parameter, ρ_R(x, p) plays the main role in the superstatistics [17–19], which describes nonequilibrium complex systems characterized by different time scales [20], and offers a route [21] to non-extensive statistical mechanics [22–25] characterized by long-tail distributions. Superstatistics has been successfully applied to broad areas of research, including hydrodynamic turbulence [26], complex networks [27], solar flares [28], high-energy physics [29], random matrix [30], and nanoscale electrochemical systems [31].

In terms of the superstatistics, our method focusing on situation (I) leads to a deterministic, time-reversible method which generates a superstatistics distribution in a dynamical manner. Generating density (8), we observe that the superstatistics distribution ρ_R(x, p) is realized by focusing only on the physical variables (x, p) in our method. Our approach is the dynamical realization of superstatistics distributions, namely, the realization of the integrand in equation (9), and is not limited to the realization of the integral in equation (9) (‘dynamical realization’ thus does not only mean realizing the superstatistics distribution ρ_R(x, p) in a certain dynamics, but also realizing the density (8), which generates ρ_R(x, p), by dynamically fluctuating β). With this realization, a physical system in a nonequilibrium environment, which is yet under control in light of the probability distribution (8), can be simulated. It must provide new gains, due to the freedom of choosing ρ_{phys} and f. Approaches to realize the marginal distribution, ρ_R(x, p), have been taken into account [32]. In other words, they utilize the distribution in static manners, for which β is not a dynamical variable, but just a variable of integration. One of the advantages of the dynamical realization of the superstatistics distribution, as compared with the static realization of the superstatistics distribution, is that we can directly observe the influence of temperature deviations on the physical system, leading to an understanding of the dynamical features of the physical system in nonequilibrium. Another advantage is that we can generate the distribution ρ_R, as well as reweighted distributions, even if the integration in equation (9), which defines ρ_R(x, p), cannot be analytically accomplished (the static method often needs the nontrivial analytical form of ρ_R). The resulting ability to freely set f guides us to direct interpretations between the simulations and the experiments. In addition, our target is a general density form of equation (8), in contrast to conventional approaches addressing a specific form, $\int d\beta e^{-U(x)} f(\beta)$, for which $\rho_{\text{phys}}(x, p, \beta) = e^{-U(x)}$. On another front, the generation of ρ_R(x, p) based on a stochastic EOM, the Langevin equation, was discussed by Beck [17], for which f(β) is assumed to exist irrespective of descriptions of the detailed structure. Hahn et al [33] characterized the EOM as stochastic differential equations driven by exchangeable processes. In our method, however, β is a realistic dynamical variable, and the specific mechanism to realize f(β) is provided. Moreover, the current deterministic method enables us to monitor the numerical accuracies in integrating the ODE [34].

Note that we shall view the variable β, employed in equation (9), as (not only the inverse temperature but also) a general parameter corresponding to issues (I)–(III) as stated above. In this sense, the term, ‘superstatistics distribution’ associated with β, will be used. It should also be noted that our aim in the present work is not to dynamically justify superstatistics, but to generate distributions compatible with the superstatistics description as defined above. Such a
justification and the relevant discussions related to it are provided in other work, such as the literature mentioned above.

In this paper, we pursue a method for generating a distribution density, equation (8), in a deterministic manner by constructing an ODE. We establish a general framework of the method and consider a mathematical structure of the resulting space of dynamical variables. In section 2 we present a vector field defining the ODE, as shown in equation (15). Then we clarify the structure of the generated probability space, to demonstrate what the realization of the probability density accurately means. In section 3 we define the total phase-space density-function, \( \rho \), suitable for our purpose, which includes the term corresponding to (8) (see equation (33)). While \( \rho \) contains other required functions that should be specifically defined, its suitable definition will lead us to the attainment of our purpose. This is ensured by theorem 6, which states the fundamental results regarding the realization of equation (8). The realization of \( \rho_r \) is established in corollary 7. In section 4 we perform the remaining task, that is, the setting of the required functions in \( \rho \), and discuss the related mathematical conditions. Note that while we can actually perform individual simulations by this procedure, we should focus on situations (I) and (II) due to the specific setting. As discussed in section 5, the numerical simulation of a simple model system demonstrates the validity of the current method, in terms of the generated distribution and the kinetics of the physical and temperature variables. We conclude our study in section 6.

2. Double density dynamics

Main idea—The target probability density, \( \rho \), of variables \( \omega \), including \( x \), \( p \), and \( Q \), where \( Q \) is a dynamical variable related to \( \beta \), can be realized by a smooth vector field \( X \) that has an invariant measure \( P \equiv \rho d\omega \) on \( \Omega \subset \mathbb{R}^N \). Namely, if we ensure a normal condition for \( \rho \), then for any solution \( \phi \) of an ODE, \( \dot{\omega} = X(\omega) \), generated by a complete field \( X \), and for any \( P \)-integrable function \( g \), we have \( \lim_{t \to \infty} \frac{1}{t} \int_0^t g(\phi(t)) dt = \int_{\Omega} g(\omega) dP(\omega) / P(\Omega) \) under an ergodic assumption [35]. Although it is not trivial to construct such a suitable \( X \) for a given \( \rho \), the density dynamics (DD) approach [8, 36], modeled on the NH method [13, 14], gives a solution for this, a kind of inverse problem. Generalizing this idea, we propose a vector field that reflects the two systems: the physical system and the parameter system. The DD for a physical system and the DD for a parameter system will be coupled in a consistent manner.

Specific strategies—To obtain \( X \) and the associated probability space we proceed as follows: (1) Construct a vector field \( X \) such that \( \text{div} \; \rho X = 0 \) holds, provided that \( \rho \) is given; (2) Establish a probability space, which is needed to accurately demonstrate the realization of the density. First, define an ODE via the obtained \( X \) along with some technical contexts. Second, assume the ergodicity. Third, calculate the long-time averages of the phase-space functions. Since we will utilize a scaled time average at this point, we should manifest the resulting probability-space structure compatible with this scaling. Accordingly we can say that \( \rho \) is realized.

2.1. Strategy (1): Construct a vector field

We begin with the definitions of the phase space \( \Omega \) and then make a general form of the vector field \( X \) defined on \( \Omega \).

Let the phase space \( \Omega \) be a domain (an open, connected set) of \( \mathbb{R}^N \) with \( N \equiv 2(n^{(1)} + n^{(2)} + 1) \), specifically given by \( \Omega \equiv D \times \mathbb{R}^{n^{(1)}} \times \mathbb{R} \times \mathbb{R}^{n^{(2)}} \times \mathbb{R}^{n^{(3)}} \times \mathbb{R} \), with \( D \) being a domain of \( \mathbb{R}^{n^{(3)}} \). For a phase-space point (or simply, a ‘state’)
\[ \omega \equiv (x^{(1)}, p^{(1)}, \zeta^{(1)}, x^{(2)}, p^{(2)}, \zeta^{(2)}) \in \Omega, \]
the first three variables \((x^{(1)}, p^{(1)}, \zeta^{(1)}) \in D \times \mathbb{R}^{d_1} \times \mathbb{R}\) describe a physical system, and the last three variables \((x^{(2)}, p^{(2)}, \zeta^{(2)}) \in \mathbb{R}^{d_2} \times \mathbb{R}^{d_3} \times \mathbb{R}\) describe a parameter system. \(x^{(a)}, p^{(a)}, \) and \(\zeta^{(a)}\) correspond to the variables \(x, p, \) and \(\zeta, \) respectively, in the NH equation \((3),\) for which finer physical descriptions will be provided in the next subsection which details the strategy \((2).\) Here, instead, we use a general description, in order to treat the variables universally and avoid any confusion. Note that, only for \(x^{(1)}\), the accessible area is defined to be not necessarily the whole but \(D,\) on which a potential function will be defined in realistic applications.

Suppose an arbitrary given density \(\rho: \Omega \to \mathbb{R},\) which is a strictly-positive and smooth \((\text{class } C^2)\) function. For this \(\rho,\) we want to define an associated \(C^1\) vector field,

\[
X: \Omega \to \mathbb{R}^N,
\]

\[
\omega \equiv \left( x^{(1)}, p^{(1)}, \zeta^{(1)}, x^{(2)}, p^{(2)}, \zeta^{(2)} \right)
\]

\[
\to \left( X_{x^{(1)}}, X_{p^{(1)}}, X_{\zeta^{(1)}}, X_{x^{(2)}}, X_{p^{(2)}}, X_{\zeta^{(2)}} \right),
\]

which satisfies the Liouville equation,

\[
\text{div } \rho X = 0.
\]

Then, due to Liouville’s theorem, these facts ensure that \(\rho d\omega\) becomes an invariant measure of the flow \(\{T_t; \Omega \to \Omega\}\) generated by the field \(X\) which is assumed to be complete (recall the framework stated in ‘Main idea’).

Although such an \(X\) is not uniquely determined, referring to the density dynamics \([8],\) we define the following vector field:

\[
X_{x^{(a)}} = h^{(a)}(\omega) \nabla \rho^{(a)} \Theta(\omega) - \nabla \rho^{(a)} h^{(a)}(\omega) \in \mathbb{R}^{d(a)},
\]

\[
X_{p^{(a)}} = - h^{(a)}(\omega) \nabla \zeta^{(a)} \Theta(\omega) + \nabla \zeta^{(a)} h^{(a)}(\omega)
- \left[ k^{(a)}(\omega) \nabla \rho^{(a)} \Theta(\omega) - \nabla \rho^{(a)} k^{(a)}(\omega) \right] p^{(a)} \in \mathbb{R}^{d(a)},
\]

\[
X_{\zeta^{(a)}} = \left[ \left( p^{(a)} \right| \nabla \rho^{(a)} \Theta(\omega) \right) - n^{(a)} \right] k^{(a)}(\omega)
- \left( \nabla \rho^{(a)} k^{(a)}(\omega) \right| p^{(a)} \right) \in \mathbb{R},
\]

with \(\cdot\) being the inner product in \(\mathbb{R}^{d(a)},\) for \(a = 1, 2,\) and

\[
\Theta \equiv -\ln \rho.
\]

Here \(h^{(a)}\) and \(k^{(a)}\) \((a = 1, 2)\) are arbitrary \(C^2\)-functions on \(\Omega,\) and they will be suitably chosen according to the problem as seen below. Namely, \(X\) is constructed from the given \(\rho\) and subsidiary functions, \(h^{(a)}\) and \(k^{(a)}\). If we consider only \(a = 1\) (that is, we only consider the physical system) and put \(h^{(1)} = k^{(1)} = 1\) (constant function), we have the original \(DD\) \([8].\)

We can confirm the smoothness \((C^2)\) of \(X\) easily and the validity of the Liouville equation via a straightforward calculation. The flow \(\{T_t\}\) generated by \(X\) becomes the basis of our method. We call this flow double density dynamics \((DDD).\)

2.2. Strategy (2): Establish a probability space

To establish a probability space defined on the ‘physical-system-space \(\times\) parameter-system-space’ and examine its structure, we will emphasize the mathematical issues involved. One can, however, consider a temperature system instead of a general parameter system to easily
obtain a physical image in one’s mind, although the issues in this subsection are applicable to a general parameter system appearing in situations (I)–(III) as demonstrated in section 1.

To proceed with the specifics, first we rewrite \( \omega \equiv (x^{(1)}, p^{(1)}, \zeta^{(1)}, x^{(2)}, p^{(2)}, \zeta^{(2)}) \) as \((x, p, \zeta, Q, \mathcal{P}, \eta)\). Here, \( \mathbb{R}^{2n} \ni (x, p) \equiv (x^{(1)}, p^{(1)}) \) are the physical variables with \( n^{(1)} \equiv n \) degrees of freedom, and \( \mathbb{R} \ni \zeta \equiv \zeta^{(1)} \) is a control variable for the physical-system temperature (here just imagine the NH mechanism for \((x, p)\) via a friction variable \( \zeta \)).

On the other hand, \((Q, P, \eta) \equiv (x^{(2)}, p^{(2)}, \zeta^{(2)})\) forms the parameter system yielding a nonequilibrium environment against the physical system, via coordinates \( Q \in \mathbb{R}^{m} \equiv \mathbb{R}^{m} \), their corresponding momenta \( P \in \mathbb{R}^{m} \), and a control variable \( \eta \in \mathbb{R}^{1} \), which again forms an analog of the NH mechanism for \((Q, P)\), as seen below. Note that, as suggested, \( \beta \) will be given as a function of \( Q \), that is, the target general ‘parameter’ \( \beta \) (which is now actually not just a constant parameter, but a dynamical variable) is treated by handling \( Q \). This is due to a technical reason, which will be discussed in section 3.1 in detail. In any case, we treat \( \beta \) in a general manner.

Next, to make an intended ODE, two technical issues are presented. Firstly, we should specifically define the subsidiary functions, and thus we use

\[
\begin{align*}
h^{(1)}(\omega) &\equiv k^{(1)}(1) \equiv 1 \text{(const. function)}, \quad (14a) \\
h^{(2)}(\omega) &\equiv k^{(2)}(1) = 1/T(x, p, Q) \quad (14b)
\end{align*}
\]

for a certain strictly-positive, \( C^2 \) function \( T \). Although equation (14) is not the unique choice of the function forms, this form may be a natural choice considering the fact that the fundamental variables we consider are \((x, p, Q)\) and that equation (14a) is utilized in the original DD, as stated above. Secondly, we multiply \( X \) by the scalar field \( T \). In general, scalar-field multiplication, that is, \( X \to \kappa X \) using a smooth, strictly-positive scalar field \( \kappa : \Omega \to \mathbb{R} \), is a general concept to re-parameterize the time for each orbit of the flow. Namely, it causes a time rescaling between a solution of the ODE \( \dot{\omega} = X(\omega) \) and the corresponding solution of \( \dot{\omega} = (\kappa X)(\omega) \). The reason why we choose \( \kappa(\omega) \equiv T(x, p, Q) \) to obtain a simple form of an EOM, as described below. This is also related to the fact that the DD also use a scalar-field multiplication with the form of \( \kappa(\omega) \equiv T = \text{const.} \) although the resulting rescaling via such a constant field is trivial, the specific effect of the current rescaling on the time averages of the phase-space functions in the DDD should be clarified, and will be investigated below.

Hence we have \( X' : \Omega \to \mathbb{R}^{n}, \omega \mapsto T(x, p, Q) X(\omega) \), represented by

\[
\begin{align*}
X'_{\omega}(\omega) &= T(x, p, Q) \nabla_{\omega} \Theta(\omega) \in \mathbb{R}^{n}, \quad (15a) \\
X'_{p}(\omega) &= -T(x, p, Q) \nabla_{p} \Theta(\omega) - T(x, p, Q) \nabla_{\omega} \Theta(\omega) p \in \mathbb{R}^{n}, \quad (15b) \\
X'_{Q}(\omega) &= T(x, p, Q) \left[ (p| \nabla_{p} \Theta(\omega)) - n \right] \in \mathbb{R}, \quad (15c) \\
X'_{\mathcal{P}}(\omega) &= \nabla_{\mathcal{P}} \Theta(\omega) \in \mathbb{R}^{m}, \quad (15d) \\
X'_{\eta}(\omega) &= -\nabla_{\eta} \Theta(\omega) - \nabla_{\omega} \Theta(\omega) \mathcal{P} \in \mathbb{R}^{m}, \quad (15e) \\
X'_{\eta}(\omega) &= (\mathcal{P} | \nabla_{\mathcal{P}} \Theta(\omega)) - m \in \mathbb{R}, \quad (15f)
\end{align*}
\]

with

\[
\bar{\Theta}(\omega) \equiv \Theta(\omega) + \ln T(x, p, Q). \quad (16)
\]
We therefore get the ODE,
\[ \dot{\omega} = X'(\omega), \] (17)
which defines the time development of the total system. This is read, e.g. for the \( x \) component, as
\[ \frac{dx}{dt} = X'_i(\omega) = T(x, p, Q) \nabla_p \Theta(\omega), \] (18)
indicating that the derivative of each coordinate with respect to time is given by \( \frac{dx}{dt} = X'_i(\omega) \)
which is equal to \( T(x, p, Q) D_p \Theta(\omega) \in \mathbb{R} \) \( (i = 1, \ldots, n) \). The ODE also determines the physical
dimension of each quantity: \( \Theta(\omega) \) is dimensionless (e.g. see equation (15c)), and so \( T(x, p, Q) \) has a
dimension of energy or temperature (e.g. see equation (18)); note that Boltzmann’s constant is unity. Note that a simple physical view of function \( T \) derived from
equations (15) and (17) is that the two systems—one the physically based system represented
by \( (x, p, \zeta) \) and the other the parameter system represented by \( (Q, \mathcal{P}, \eta) \)—interact through
(as well as \( \rho(\omega) \) ) a quantity \( T(x, p, Q) \), which depends only on the fundamental variables
\( (x, p, Q) \) and plays the role of energy or temperature as implied by its physical dimension.

The next task is to show that the obtained ODE (17) produces an individual orbit that has
a required property: namely, by assuming ergodicity, a time average equals the space average
for any phase-space function \( g \). Here, the time average is not the simple one (which is denoted
by \( g' \) ), but a scaled one, due to the scalar-field multiplication adopted above. We shall show
\[ \bar{g} = \langle g \rangle, \] (19)
where the lhs of equation (19) is a scaled long-time average of \( g \), defined by
\[ \bar{g} \equiv \frac{\bar{g}T}{T}, \] (20)
and the rhs of equation (19) is the space average of \( g \) weighted by \( \rho \) and written as
\[ \langle g \rangle \equiv \int_{\Omega} g(\omega) \rho(\omega) d\omega = \int_{\Omega} \rho(\omega) d\omega. \] (21)
We prove equation (19) by proposition 1 and then clarify the resulting probability space
compatible with the scaled time average. Before doing so, to accurately state these and
subsequent issues we shall use the following definitions.

Definitions and notations. We define projections: \( \pi_1: \Omega \to \Omega_1 := D \times \mathbb{R}^n \times \mathbb{R}^m, \)
\( \omega \equiv (x, p, \zeta, Q, \mathcal{P}, \eta) \mapsto (x, p, Q) \); \( \pi_2: \Omega \to D \times \mathbb{R}^n, \omega \mapsto (x, p) \); similarly, \( \pi_3(\omega) := Q \)
and \( \pi_4(\omega) := x \). To clarify a variable dependence, we denote \( \overline{T} \equiv T_{\pi_1}; \Omega \to \mathbb{R}, \)
\( \omega \mapsto T(x, p, Q) \). For any \( n \in \mathbb{N}, l_n \) and \( \mathcal{L}_n \) represent the Lebesgue measure and the Lebesgue
measurable sets on \( \mathbb{R}^n \), respectively; \( \mathcal{B}_n \) represents the Borel measurable sets on \( \mathbb{R}^n \); \( \mathcal{L}_n \cap \Omega \)
is denoted by \( \mathcal{L}_n^{\pi_1} \) and so on. Basically, we consider a measure space \( (\Omega, \mathcal{L}_n^{\pi_1}, P) \), where
\( P = \rho d\mathcal{L}_n \), or employ a measure \( P' = \rho' d\mathcal{L}_n \) using a modified density \( \rho' \) instead of the given
density \( \rho \), as described below. Put \( L^1(\rho) \equiv \{ g: \Omega \to \mathbb{R} \} g \) is \( \mathcal{B}_n \)-measurable and \( P \)-integrable.
For any \( n \in \mathbb{N} \) and \( B \subset \mathbb{R}^n, \chi_B^\omega(\omega) \) is the characteristic function of \( B \) (that is, \( \chi_B^\omega(\omega) = 1 \)
if \( \omega \in B \) and \( \chi_B^\omega(\omega) = 0 \) otherwise), but we may omit \( n \) if it is clear. \( \mathbb{R}_+ \) denotes the strictly-positive real numbers.
Equation (19) is valid as follows:

**Proposition 1.** Let \( \int_{\Omega} \rho' d\mathcal{L}_n < \infty \), where \( \rho' \equiv \rho/\overline{T} \). Assume \( X' = \overline{T} X \) is complete, so that
its flow \( \{ T_t': \Omega \to \Omega \} \) is generated. Then for \( g \in L^1(P) \), the following statements hold:
(i) for a.e. \( \omega \in \Omega \) there exists a limit,

\[
\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau (g \cdot \dot{T})(T'_t(\omega))dt = \overline{gT}(\omega) \in \mathbb{R};
\]  

(22)

(ii) if \( \{T'_t\} \) is ergodic with respect to the measure space \((\Omega, \mathcal{L}^\mathbb{R}_N, P')\), then

\[
\overline{gT}(\omega) = \int_\Omega g\rho dl_N \left/ \int_\Omega \rho' dl_N \right. \in \mathbb{R} \quad \text{for a.e. } \omega \in \Omega;
\]

(23)

(iii) furthermore, if \( \rho \) is \( l_N \)-integrable, then

\[
\overline{gT}(\omega) = \int_\Omega g\rho dl_N \left/ \int_\Omega \rho dl_N \right. \in \mathbb{R} \quad \text{for a.e. } \omega.
\]

(24)

**Proof.** Since \( \rho' \colon \Omega \to \mathbb{R} \) becomes normal, i.e. smooth \((C^2)\), strictly positive, and \( l_N \)-integrable, and since \( \Omega \) is open, \( P' = \rho' dl_N \) becomes a strictly-positive, finite measure on \( \mathcal{L}^\mathbb{R}_N \). For \( X' \), it becomes \( C^1 \), and the Liouville equations holds:

\[
\text{div } \rho'X' = \text{div } \rho X = 0.
\]

(25)

From these facts and the assumption of the completeness of \( X' \), \( P' \) becomes an invariant measure with respect to \( \{T'_t\} \), i.e.

\[
\text{with } \forall t \in \mathbb{R}, \forall A \in \mathcal{L}^\mathbb{R}_N, P'(T^{-1}_t A) = P'(A).
\]

Now for a function \( g \in L^1(P) \), it holds that \( g' \equiv \overline{gT} \in L^1(P') \) due to

\[
\mathbb{R} \ni \int_\Omega gdP = \int_\Omega g\rho dl_N = \int_\Omega g'\rho' dl_N = \int_\Omega g'dP',
\]

(26)

and \( g'\Psi \colon \mathbb{R} \times \Omega \to \mathbb{R} \) is \( B_1 \otimes \mathcal{L}^\mathbb{R}_N \)-measurable, where \( \Psi(t, \omega) \equiv T'_t(\omega) \) for \( (t, \omega) \in \mathbb{R} \times \Omega \). (i) Thus, according to Birkhoff’s ergodic theorem, for \( P'\)-a.e. \( \omega \in \Omega \) there exists a time average of \( g' \), i.e. \( g' \equiv \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau g'(T'_t(\omega))dt \in \mathbb{R} \), which is equivalent to \( \overline{gT}(\omega) \) defined in equation (22) (we abbreviate as \( \overline{gT} = \overline{gT} \equiv \overline{gT} \)). Note ‘\( P'\)-a.e.’ can be replaced by \( P\)-a.e. or \( l_N\)-a.e., since \( P' \), \( P \), and \( l_N \) become equivalent (i.e. absolutely continuous with each other) because \( \rho, \rho' > 0 \), so that we abbreviate them to just ‘a.e.’ (ii) it follows from the ergodicity that \( \overline{gT}(\omega) = \int_\Omega g'dP' / \rho' dl_N \) for \( P'\)-a.e. \( \omega \), where the rhs equals \( \int_\Omega g' dl_N / \int_\Omega \rho' dl_N \) due to equation (26). (iii) Since the \( l_N \)-integrability of \( \rho \) is nothing but the \( P \)-integrability of \( g \equiv 1 \) (constant function), the above conditions (i) and (ii) apply to \( 1 \in L^1(P) \), implying

\[
\overline{T}(\omega) \equiv \exists \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau T(T'_t(\omega))dt = \int_\Omega \rho dl_N / \int_\Omega \rho' dl_N \in \mathbb{R} \quad \text{(a.e. } \omega) \quad \text{(27)}
\]

Also note that \( \overline{T}(\omega) = 0 \) (a.e.) since \( \int_\Omega \rho dl_N > 0 \). Thus, the division of equation (23) by equation (27) leads to the validity of equation (24) for a.e. \( \omega \). \( \square \)

Hence, the accurate description of equation (19) is given in the context of equation (24), which now becomes
In the current method, we thus realize the probability measure \( \hat{P} \equiv P / P(\Omega) \) and the probability space \((\Omega, \mathcal{U}_N, \hat{P})\), under the assumption of the ergodicity and the integrabilities of \( \rho \) and \( \rho' \). This is done through the ability of calculating the expectation value, \( \int_\Omega g d\hat{P} / P(\Omega) \), for any function \( g \in L^1(P) \) via the scaled-time average \( \overline{g} \). Namely, as is explicitly stated, the probability for any set \( B \in \mathcal{B}_N^D \) can be defined and represented as \[ \nu(B) = \overline{\chi_B} = \left\langle \chi_B^N \right\rangle = \hat{P}(B), \] (30)
which implies that we get \( \Omega, B_N^\Omega, \nu \) and \( \nu = \hat{P}|B_N^\Omega \) (strictly speaking, the equality \( \overline{\chi_B} = \left\langle \chi_B^N \right\rangle \) in equation (30) is valid for the \( l_N \)-a.e. initial point \( \omega \in \Omega \), and the set of all such points should be written as \( \Omega_B \) considering the dependence on uncountably many \( B \). Thus it should be \( \nu(B) = \overline{\chi_B}(\omega_B) \) for any \( \omega_B \in \Omega_B \). Similar notation applies hereafter. Therefore, using the equivalence of \( \hat{P} \) and \( l_N \), the completion of \((\Omega, \mathcal{U}_N^D, \hat{P})\) emerges as the whole probability space \((\Omega, \mathcal{U}_N^D, \hat{P})\).

Remark 2. We see that proposition 1 holds for any \( C^2 \), strictly-positive \( \kappa : \Omega \to \mathbb{R} \), instead of \( \hat{T} \). This means that the probability-space description in a usual ergodic dynamical system, which is generated by a vector field having a smooth invariant density \( \rho \), applies for a dynamical system that is described by a scalar-field-multiplied vector-field along with keeping the density \( \rho \), as long as we consider the scaled-time average. Of course \( \kappa = 1 \) reduces to the usual description. Thus we have considered a general situation to obtain a probability-space structure.

Example 3. A marginal distribution for the \( x \)-component is given by
\[ B_n^D \to \mathbb{R}, \quad B_1 \mapsto \overline{\chi_{B_1}^N \circ \pi_x}, \] (31)
as confirmed from \( \overline{\pi_x} \circ \pi_x = \chi_{B_1 \times \mathbb{R}^m}^N \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \) and equation (30), indicating
\[ \overline{\chi_{B_1}^N \circ \pi_x} = \hat{P}(B_1 \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}) (\forall B_1 \in B_n^D), \] (32)
whose rhs implies the marginal distribution. Marginal distributions for other components are given similarly.

3. Realization of a joint distribution of physical and parameter systems

In section 2.2, the phase-space density, \( \rho \), has been treated in a very general manner, where we have only assumed the following mathematical conditions: smoothness, positivity, and integrability. To perform our purpose, namely, to create the joint distribution density, equation (8), we should set a suitable specific form of \( \rho \) along with satisfying the three conditions. This is established in the subsequent section 3.1. The fundamental results, including the creation of the joint distribution density, will be demonstrated in section 3.2.
Regarding the function $T$, in this section it still retains its abstract form, simply with the condition of the smoothness and positivity, and we will discuss in section 4 its more specific form.

3.1. Density form

To create density (8), we set as follows:

$$\rho: \Omega \to \mathbb{R}, \omega \mapsto \rho_{\text{Phys}}(x, p, \sigma(Q)) \rho_f(Q) \rho_{\text{Cntt}}(\mathcal{P}, \zeta, \eta),$$

(33)

$$\rho_f: \mathbb{R}^m \to \mathbb{R}, Q \mapsto f(\sigma(Q)) |\det D\sigma(Q)|.$$  

(34)

Each function constituting $\rho$ and function $T$ have to satisfy the following six conditions:

(C1) $\sigma: \mathbb{R}^m \to \mathbb{R}^m$ is a $C^3$-diffeomorphism, and $\det D\sigma: \mathbb{R}^m \to \mathbb{R}$ has a definite signature.

Note that $\sigma$ is introduced to control the admissible space of $\beta$ via mapping $Q \mapsto \sigma(Q)$ that moves in a whole free space $\mathbb{R}^m$, which is convenient to establish a flow. $|\det D\sigma(Q)|$ in equation (34) appears as the Jacobian deduced from this mapping, and it is unity if one uses mapping $\sigma$ that is the identity, where $Q = \beta$.

(C2) $\rho_{\text{Cntt}}: \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, is of class $C^2$ and $l_{m+2}$-integrable, that is, $Z_{\text{Cntt}} \equiv \int_{\mathbb{R}^m} \rho_{\text{Cntt}}(\mathcal{P}, \zeta, \eta) d\mathcal{I}_{m+2}(\mathcal{P}, \zeta, \eta) \in \mathbb{R}$.

Note that the fundamental variables to perform our purpose are the physical variables $(x, p)$ and the parameter variable $Q$, for which their density form should be fixed in view of equation (8). Thus we choose a simple and general expression of the density form regarding the remaining variables $\mathcal{P}$, $\zeta$ and $\eta$, which seem to be subsidiary but control the fundamental variables in total.

(C3) $\rho_{\text{Phys}}: D \times \mathbb{R}^m \times \sigma(\mathbb{R}^m) \to \mathbb{R}_+, (x, p, \beta) \mapsto \rho_{\text{Phys}}(x, p, \beta)$, is of class $C^2$.

(C4) $f: \sigma(\mathbb{R}^m) \to \mathbb{R}_+$ is (i) of class $C^2$ and $l_m$-integrable.

Note that (C4ii) seems to be a natural condition, but will be required only for ensuring equation (55) in the following contexts.

(C5) $D \times \mathbb{R}^m \times \sigma(\mathbb{R}^m) \to \mathbb{R}_+, (x, p, \beta) \mapsto \rho_{\text{Phys}}(x, p, \beta)f(\beta)$, is $l_{2m+1}$-integrable.

(C6) $T: D \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}_+, (x, p, Q) \mapsto T(x, p, Q)$, is (i) of class $C^2$; and (ii) $D \times \mathbb{R}^m \times \sigma(\mathbb{R}^m) \to \mathbb{R}_+, (x, p, \beta) \mapsto \rho_{\text{Phys}}(x, p, \beta)f(\beta)/T(x, p, \sigma^{-1}(\beta))$, is $l_{2m+1}$-integrable.

Although (C6i) has already been assumed, we state it here for clarity.

We show that $\rho$ defined above satisfies the required properties:

**Proposition 4.** Under conditions (C1)–(C4i) and (C5), $\rho$ becomes smooth ($C^2$), strictly positive, and $l_N$-integrable such that

$$\int_{\Omega} \rho dV = \int_{D \times \mathbb{R}^m \times \sigma(\mathbb{R}^m)} \rho_{\text{Phys}}(x, p, \beta)f(\beta) d\mathcal{I}_{2m+1}(x, p, \beta) \times Z_{\text{Cntt}} \in \mathbb{R}_+$$

(35a)

$$= \int_{D \times \mathbb{R}^m} \rho_{\text{Phys}}(x, p) d\mathcal{I}_{2m}(x, p) \times Z_{\text{Cntt}}$$

(35b)
\[ = \int_{\Omega} f(\beta)Z(\beta)d\lambda_n(\beta) \times Z_{\text{Cut}}, \tag{35c} \]

where
\[ \rho_R: D \times \mathbb{R}^n \to \mathbb{R}, \quad (x, p) \mapsto \int_{\sigma(\mathbb{R}^n)} \rho_{\text{Phys}}(x, p, \beta) f(\beta) \, d\lambda_n(\beta), \tag{36} \]

and
\[ Z: \sigma(\mathbb{R}^m) \to \mathbb{R}, \quad \beta \mapsto \int_{D \times \mathbb{R}^n} \rho_{\text{Phys}}(x, p, \beta) d\lambda_2(x, p), \tag{37} \]

are Borel measurable, strictly positive, and almost everywhere finite.

**Proof.** The smoothness and positivity of \( \rho \) are clear from the conditions (C1)–(C4i) (C1) implies that \( \mathbb{R}^m \ni Q \mapsto |\det D\sigma(Q)| = \pm \det D\sigma(Q) \) is \( C^2 \) and nonzero, where either sign is fixed. \( \int_{\Omega} \rho N > 0 \) is thus also clear. We begin with splitting the phase space \( \Omega \) into \( \Omega_1 = D \times \mathbb{R}^n \times \mathbb{R}^m \), to which the ‘fundamental’ variables \((x, p, Q)\) belong, and \( \Omega_2 := \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \), to which the ‘control’ variables \((P, \zeta, \eta)\) belong. Although \( \Omega = \Omega_1 \times \Omega_2 \) due to our definition, by using a variable-order exchange, we have
\[ \int_{\Omega} \rho \, dl_N = \int_{\Omega_1} \rho_1 \, dl_{2n+m} \cdot \int_{\Omega_2} \rho_{\text{Cut}} \, dl_{m+2}, \tag{38} \]

where
\[ \rho_1: \Omega_1 \to \mathbb{R}, \quad (x, p, Q) \mapsto \rho_{\text{Phys}}(x, p, \sigma(Q)) \rho_f(Q), \tag{39} \]

and \( \rho_{\text{Cut}}: \Omega_2 \to \mathbb{R} \) are strictly-positive and \( C^2 \) functions. The integration of \( \rho_{\text{Cut}} \) in equation (38) is \( Z_{\text{Cut}} \) (see (C2)), and the integration of \( \rho_1 \) in equation (38) will be examined below. It follows from Fubini’s theorem being applied to the positive, \( \mathcal{B}_{2n+m}(=\mathcal{B}_{2n} \otimes \mathcal{B}_m) \) -measurable function \( \rho_1 \) that
\[ \int_{\Omega_1} \rho_1 \, dl_{2n+m} \]
\[ = \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \int_{\mathbb{R}^m} \rho_1(x, p, Q) \, d\lambda_m(Q). \tag{40a} \]

where \( g^{(x,p)}: \sigma(\mathbb{R}^m) \to \mathbb{R}^+, \quad \beta \mapsto \rho_{\text{Phys}}(x, p, \beta)f(\beta), \) that is, \( \rho_1(x, p, Q) = g^{(x,p)}(\sigma(Q)) |\det D\sigma(Q)| \, d\lambda_m(Q). \) For all \((x, p) \in D \times \mathbb{R}^n,\) \( g^{(x,p)} \) is positive and \( L^m(\mathbb{R}^m) \)-measurable, so that a variable transformation formula being applied to diffeomorphic \( \sigma \) leads to
\[ \int_{\mathbb{R}^m} g^{(x,p)}(\sigma(Q)) |\det D\sigma(Q)| \, d\lambda_m(Q) = \int_{\sigma(\mathbb{R}^m)} g^{(x,p)}(\beta) \, d\lambda_m(\beta) = \rho_R(x, p). \tag{41} \]

Thus
\[ \int_{\Omega_1} \rho_1 \, dl_{2n+m} = \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_R(x, p), \tag{42a} \]
\[
\int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \int_{\sigma(\mathbb{R}^n)} dl_m(\beta) \rho_{\text{phys}}(x, p, \beta) f(\beta) = \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{\text{phys}}(x, p, \beta) f(\beta)
\]

(42b)

\[
\int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{\text{phys}}(x, p, \beta) f(\beta) dl_{2n+m}(x, p, \beta) \in \mathbb{R}
\]

(42c)

\[
\int_{\sigma(\mathbb{R}^n)} dl_m(\beta) \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{\text{phys}}(x, p, \beta) f(\beta)
\]

(42d)

where we have applied Fubini’s theorem to the positive, \(B_{2n+m}^{D \times \mathbb{R}^n} = B_{2n}^{D \times \mathbb{R}^n} \otimes B_{m}^{(\mathbb{R}^n)}\) -measurable function, \((D \times \mathbb{R}^n) \times \sigma(\mathbb{R}^m) \rightarrow \mathbb{R}, (x, p, \beta) \mapsto \rho_{\text{phys}}(x, p, \beta) f(\beta)\), and used the integrability condition (C5) for the finiteness in equation (42c). Therefore we obtain the integrability of \(\rho\) and the integral formulae in equations (35a) and (35b).

The measurability of \(\rho_R\) follows from the fact that \(D \times \mathbb{R}^n \rightarrow \mathbb{R}, (x, p) \mapsto \int_{\mathbb{R}^n} \rho_1(x, p, Q) dl_m(Q)\) becomes \(\mathcal{B}_{2n}^{D \times \mathbb{R}^n}\)-measurable (Fubini’s theorem in equation (40)) and the fact that this map is \(\rho_R\) (equation (41)). The positivity of \(\rho_R\) is evident. \(\rho_R\) is finite for \(l_{2n}\)-a.e., due to the integrability, indicated in equation (42c), through Fubini’s theorem. Similarly, we see that \(Z': \sigma(\mathbb{R}^n) \rightarrow \mathbb{R}, \beta \mapsto \int_{D \times \mathbb{R}^n} \rho_{\text{phys}}(x, p, \beta) f(\beta) dl_{2n}(x, p)\) is well defined, \(\mathcal{B}_{\sigma(\mathbb{R}^m)}^{\mathbb{R}^n}\)-measurable, finite for \(l_m\)-a.e. \(\beta\), and strictly positive. So does \(Z'/f\), which equals \(Z\) (note \(0 < f < \infty\)). Finally, observe that equation (42d) becomes \(\int_{\sigma(\mathbb{R}^n)} dl_m(\beta) Z'(\beta) = \int_{\sigma(\mathbb{R}^n)} f(\beta) Z(\beta) dl_m(\beta)\), which leads to equation (35c).

\[\square\]

As found from the above proof, without the integrability condition (C5), both \(\rho_R\) and \(Z\) are well defined and Borel measurable, and (C5) is equivalent to both the finiteness of the integral in equations (35b) and that in equation (35c).

### 3.2. Fundamental results

The dynamics for flow \(\{T_t\}\), generated by the vector field \(X\) and the scalar function \(T\), and the associated probability space \((\Omega, \mathcal{L}^2, \mu)\) were formulated through proposition 1, where \(X\) was constructed from an abstract form of density function \(\rho\) which should satisfy just the normal three conditions. These conditions are satisfied for \(\rho\) which is specifically defined by equations (33) and (34) under (a part of) conditions (C1)–(C5), as shown in proposition 4.

In the following text, \(\rho\) is the density defined by equations (33) and (34), and \(\rho' = \rho/\tilde{T}\). In this subsection, we first prove lemma 5 to show that the integrability of \(\rho'\) required in proposition 1 is ensured by (C6). Next, we explicitly show that by using \(\rho\) we can realize the joint distribution density (8), that is, we can dynamically realize the superstatistics distribution. This is stated in theorem 6, as a special case of equation (24) with \(g(\omega) \equiv B(x, p, \sigma(Q))\), a function of fundamental variables \((x, p, Q)\). Integral formulae for \(B\) are also shown in lemma 5.

Consider the condition:

(C7) (i) \(\Omega_1^T : D \times \mathbb{R}^n \times \sigma(\mathbb{R}^m) \rightarrow \mathbb{R}, (x, p, \beta) \mapsto B(x, p, \beta)\), is \(B_{2n+m}^{D \times \mathbb{R}^n}\)-measurable, and (ii) a function \(\Omega_1^T \rightarrow \mathbb{R}, (x, p, \beta) \mapsto B(x, p, \beta) \rho_{\text{phys}}(x, p, \beta) f(\beta)\), is \(l_{2n+m}\)-integrable.

Note, instead of \(B\), we can treat \(b: \Omega_1 = D \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}, (x, p, Q) \mapsto b(x, p, Q)\), since \(B\) and \(b\) can be defined from the one to the other in a trivial manner such that \(b(x, p, Q) = B(x, p, \sigma(Q))\) and \(B(x, p, \beta) = b(x, p, \sigma^{-1}(\beta))\).
Lemma 5. Under conditions (C1)–(C4i), (C5), and (C7i), condition (C7ii) is equivalent to
\[ \int_{\Omega} b(x, p, Q) \rho(\omega) d\lambda_\omega(\omega) \in \mathbb{R}. \]
Under (C1)–(C4i), (C5) and (C7), we have
\[ \mathbb{R} \ni \langle b \circ \eta \rangle \]
\[ = \int_{\Omega} B(x, p, \beta) \rho_{\text{phys}}(x, p, \beta) f(\beta) d\lambda_{2n+m}(x, p, \beta) \]
\[ = \int_{D \times \mathbb{R}^n} d\lambda_{2n+m}(x, p) \int_{\sigma(\mathbb{R}^m)} dl_m(\beta) B(x, p, \beta) \rho_{\text{phys}}(x, p, \beta) f(\beta) \]
\[ = \int_{\sigma(\mathbb{R}^m)} f(\beta) Z(\beta) dl_m(\beta). \]

Proof. Assume (C1)–(C4i), (C5), and (C7i). Obviously \( b \circ \eta \) is \( B_{\infty}^1 \)-measurable. Below we shall calculate \( \langle b \circ \eta \rangle \), for which similar procedures to the proof of proposition 4 will be used with careful attention to the indefinite signature of \( B \). We have
\[ \int_{\Omega} \left| b \circ \eta \right| \rho d\lambda_N = \int_{\Omega} \left| \rho' \right| d\lambda_{2n+m} \cdot Z_{\text{Conr}}, \]
with \( \rho'_i = b \rho_i ; \Omega_i \rightarrow \mathbb{R} \) (see equation (39)), noting that this map is equal to \( (x, p, Q) \mapsto \tilde{g}^{(x,p)}(\sigma(Q)) \det D\sigma(Q) \), where \( B_{\infty}^1(\mathbb{R}^m) \)-measurable function \( \tilde{g}^{(x,p)}: \sigma(\mathbb{R}^m) \rightarrow \mathbb{R} \), \( \beta \mapsto (B \rho_{\text{phys}})(x, p, \beta) f(\beta) \), is defined for every \( (x, p) \). Thus, a variable transformation formula and Fubini’s theorem read
\[ \int_{\Omega} \left| \rho'_i \right| d\lambda_{2n+m} = \int_{D \times \mathbb{R}^n} d\lambda_{2n+m}(x, p) \int_{\sigma(\mathbb{R}^m)} dl_m(\beta) \left( B \rho_{\text{phys}} \right)(x, p, \beta) f(\beta) \]
\[ = \int_{D \times \mathbb{R}^n} \left| (B \rho_{\text{phys}})(x, p, \beta) \rho_{\text{phys}}(x, p, \beta) f(\beta) dl_{2n+m}(x, p, \beta). \]

Also assume (C7ii). Then we have \( \int_{\Omega} \left| \rho'_i \right| d\lambda_{2n+m} \in \mathbb{R} \), implying \( b \circ \eta \in L^1(P) \). It follows from this integrability that \( \int_{\Omega} b \circ \eta \rho d\lambda_N = \int_{\Omega} \rho'_i d\lambda_{2n+m} \cdot Z_{\text{Conr}} \)
and
\[ \mathbb{R} \ni \int_{\Omega} \rho'_i d\lambda_{2n+m} = \int_{D \times \mathbb{R}^n} d\lambda_{2n+m}(x, p) \int_{\mathbb{R}^m} dl_m(Q) \rho'_i(x, p, Q). \]

Here, for \( L_{2n} \text{a.e. } (x, p), \rho'_1, (x, p), \mathbb{R}^m \rightarrow \mathbb{R}, \sigma(i) \mapsto \rho'_i(x, p, Q) \), becomes \( L_{m} \)-integrable, that is,
\[ \mathbb{R} \ni \int_{\mathbb{R}^m} \rho'_1(Q) d\lambda_m(Q) = \int_{\mathbb{R}^m} \tilde{g}^{(x,p)}(\sigma(Q)) \det D\sigma(Q) dl_m(Q). \]
\[ \mathbb{R} \ni \int_{\mathbb{R}^n} \rho_1 \left( x, p, \mathbf{Q} \right) d\mathbf{m}(\mathbf{Q}) = \int_{\sigma(\mathbb{R}^r)} \left( B\rho_{\text{phys}} \right)(x, p, \beta)f(\beta)d\mathbf{m}(\beta) \text{ a.e.} (x, p). \]

Condition (C7ii) thus leads us to
\[ \mathbb{R} \ni \int_{\Omega} \rho_1' d\mathbf{m}_{2n+m} \]
\[ = \int_{\mathbb{D} \times \mathbb{R}^r} d\mathbf{m}_{2n}(x, p) \int_{\sigma(\mathbb{R}^r)} \left( B\rho_{\text{phys}} \right)(x, p, \beta)f(\beta)d\mathbf{m}(\beta) \]
\[ = \int_{\mathbb{D} \times \mathbb{R}^r \times \sigma(\mathbb{R}^r)} \left( B\rho_{\text{phys}} \right)(x, p, \beta)f(\beta)d\mathbf{m}_{2n+m}(x, p, \beta) \]
\[ = \int_{\sigma(\mathbb{R}^r)} \mathbf{m}(\beta)f(\beta) \int_{\mathbb{D} \times \mathbb{R}^r} d\mathbf{m}_{2n}(x, p) \left( B\rho_{\text{phys}} \right)(x, p, \beta). \]

Note \( \int_{\Omega} \rho d\mathbf{m}_N \in \mathbb{R}_+ \) by proposition 4. From equation (46), the fact that \( \langle b \circ \pi \rangle = \int_{\Omega} b \circ \pi \cdot \rho d\mathbf{m}_N / \int_{\Omega} \rho d\mathbf{m}_N = \int_{\Omega} \rho_1' d\mathbf{m}_{2n+m} \cdot Z_{\text{Cov}} / \int_{\Omega} \rho d\mathbf{m}_N \in \mathbb{R}, \) and equation (35), we see that \( \langle b \circ \pi \rangle \) equals equations (43b)–(43d).

The above results have stated that \( (C7ii) \implies \int_{\Omega} b \circ \pi \cdot \rho d\mathbf{m}_N \in \mathbb{R}, \) under the conditions (C1)–(C4i), (C5), and (C7i). Conversely, if \( \int_{\Omega} b \circ \pi \cdot \rho d\mathbf{m}_N \in \mathbb{R} \) is assumed, then equations (44) and (45) manifest (C7ii) [note \( Z_{\text{Cov}} > 0 \) due to (C2)].

**Theorem 6.** (i) Let \( \rho \) and \( T \) satisfy (C1)–(C6) (except (C4ii)); and (ii) let \( X' \equiv T X \) be complete and its flow \( \{ T_t^1; \Omega \to \Omega | t \in \mathbb{R} \} \) be ergodic with respect to the measure space \( (\Omega, L_*^1, P') \). Then for \( B \) satisfying (C7) we have, for a.e. \( \omega, \)
\[ \mathbb{R} \ni b \circ \pi(\omega) = \langle b \circ \pi \rangle, \]
which is also represented by equations (43b)–(43d).

**Proof.** From (C1)–(C4i), and (C5), we can apply proposition 4, so that \( \rho \) satisfies the three conditions. Due to this fact and (C6i), \( P = \rho d\mathbf{m}_N \) and \( P' = \rho' d\mathbf{m}_N \) are well-defined measures on \( (\Omega, L_*^1) \) in the course of applying proposition 1. Lemma 5 states \( b \circ \pi \in L^1(P) \). In addition, by temporarily putting \( b \equiv 1/T, \) the corresponding \( B \) satisfies (C7) due to the condition of \( T, \) that is, (C6), so that lemma 5 also states \( \mathbb{R} \ni \int_{\Omega} b(x, p, \mathbf{Q}) \rho(\omega)d\mathbf{m}_N(\omega) = \int_{\Omega} \rho' d\mathbf{m}_N. \) From these results and assumption (ii), proposition 1 with \( g \equiv b \circ \pi \) (this \( b \) is the one in equation (47)) demonstrates \( \exists \beta \circ \pi(\omega) = \langle b \circ \pi \rangle \in \mathbb{R} \) for a.e. \( \omega. \) Lemma 5 shows the equality among \( \langle b \circ \pi \rangle \) and equations (43b)–(43d).

In this subsection we shall assume that the conditions (i) and (ii) in theorem 6 hold, so that propositions 1 and 4 are also applicable (\( \int_{\Omega} \rho' d\mathbf{m}_N < \infty \) is concluded) as shown in the proof of theorem 6.

Theorem 6 shows that the joint distribution density, equation (8), can be created by the scaled long-time average generated by the flow \( \{ T_t^1 \}. \) This is explicitly stated as follows. For any \( M \in \mathbb{B}_{2n+m}, B \equiv B_M \equiv \chi_{M2n+m} \) satisfies (C7), so that theorem 6 demonstrates
\[ \mathbb{R} \ni b_M \circ \eta_1(\omega) \]  
\[ = \int_{\Omega} \chi_M^{2n+m}(x, p, \beta) \rho_{\text{phys}}(x, p, \beta)f(\beta) d\mathbb{L}^{2n+m}(x, p, \beta) \frac{\rho_{\text{phys}}(x, p, \beta)}{\rho_{\text{phys}}(x, p, \beta)f(\beta) d\mathbb{L}^{2n+m}} \hspace{1cm} (\text{a.e. } \omega) \]  
\[ = \int_M \rho_0(x, p, \beta) d\mathbb{L}^{2n+m}(x, p, \beta), \]  
\[ \text{where} \]  
\[ \rho_0(x, p, \beta) = \frac{\rho_{\text{phys}}(x, p, \beta)}{\rho_{\text{phys}}(x, p, \beta)f(\beta) d\mathbb{L}^{2n+m}(x, p, \beta),} \]  
and \( b_M(x, p, Q) \) takes 1 if \((x, p, \sigma(Q)) \in M\) and otherwise it takes 0. Thus we obtain, via the scaled long-time average, a specific probability measure (distribution),  
\[ \nu_B: B_{2n+m} \rightarrow \mathbb{R}, \quad M \xrightarrow{\text{d}} b_M \circ \eta_1 = \int_M \rho_0 d\mathbb{L}^{2n+m}, \]  
whose Radon–Nikodym derivative with respect to \( \mathbb{L}^{2n+m} \) is \( \rho_0 \) for \( \mathbb{L}^{2n+m} \)-a.e. (note \( \rho_0 \) is positive and integrable), that is, \( \frac{\rho_0}{\mathbb{L}^{2n+m}}(x, p, \beta) = \rho_0(x, p, \beta) \) (a.e.), which is proportional to \( \rho_{\text{phys}}(x, p, \beta)f(\beta) \). This fact indicates the realization of density (8).

The following indicates the realization of the superstatistics distribution.

**Corollary 7.** For \( A(x, p) \), a function of pure physical variables such that
\[ (C8) \quad A: D \times \mathbb{R}^n \rightarrow \mathbb{R} \] is \( B_{2n}^D \times \mathbb{R}^n \)-measurable and \( \rho_K d\mathbb{L}^{2n} \)-integrable,
we have
\[ \overline{A} \circ \overline{\rho_K}(\omega) = \left( A \right)_{\rho_K} := \frac{\int_{D \times \mathbb{R}^n} A(x, p) \rho_K(x, p) d\mathbb{L}^{2n}(x, p)}{\int_{D \times \mathbb{R}^n} \rho_K d\mathbb{L}^{2n}} \in \mathbb{R} \] \hspace{1cm} (a.e. } \omega. \]  
\[ \text{Here the integrability condition in (C8) can be replaced by a condition such that } \Omega'_1 \rightarrow \mathbb{R}, \]  
\((x, p, \beta) \mapsto A(x, p) \rho_{\text{phys}}(x, p, \beta)f(\beta), \) is \( \mathbb{L}^{2n+m} \)-integrable.

**Proof.** Note the two integrability conditions are equivalent:
\[ \int_{\Omega} |A(x, p)| \rho_{\text{phys}}(x, p, \beta)f(\beta) d\mathbb{L}^{2n+m}(x, p, \beta) \]  
\[ \frac{\int_{D \times \mathbb{R}^n} A(x, p) \rho_K(x, p) d\mathbb{L}^{2n}(x, p)}{\int_{D \times \mathbb{R}^n} \rho_K d\mathbb{L}^{2n}} \]  
\[ = \int_{D \times \mathbb{R}^n} |A(x, p)| \rho_{\text{phys}}(x, p, \beta)f(\beta) d\mathbb{L}^{2n}(x, p, \beta) \]  
\[ = \int_{D \times \mathbb{R}^n} |A(x, p)| \rho_K(x, p). \]  
Thus, if we define \( B_X: \Omega'_1 \rightarrow \mathbb{R}, \]  
\((x, p, \beta) \mapsto A(x, p) \), then \( B_X \) meets (C7). Applying theorem 6 to \( B \equiv B_A \), where \( b_\eta \circ \eta_1 = A \circ \eta_1 \), we get
Equation (51) states that superstatistics distribution \( \rho_R \) is realized by calculating the scaled long-time average of the physical quantities. It should be noted that for realizing \( \rho_R \) the current scheme needs no explicit form of \( \rho_R \), which is defined by the integration formula (36). It is an advantage of the dynamical realization of superstatistics that \( \rho_R \) is realized regardless of whether the integration can be explicitly done or not, as stated in section 1.

Note on ergodicity. So far, we have assumed the ergodicity with respect to \( \{ T_t \} \) and \((\Omega, L^0, P^\prime)\). Proof of the ergodicity is hard to achieve in general, and we here just note that the following two obstacles \([8]\) of the ergodicity vanish in the current system: fixed points (i.e. \( \omega_0 \in \Omega \) such that \( X(\omega_0) = 0 \)) and zero divergence (i.e. \( \text{div} X'(\omega) = 0 \) for \( \forall \omega \in \Omega \)). For the first issue, we can easily observe from equations (15) and (15) and \( m > 0 \) that the fixed points do not exist. For the second issue, we have

\[
\text{div} X'(\omega) = \left( \nabla_x T(x, p, Q) \right| \nabla_p \Theta(\omega)) - \left( \nabla_p T(x, p, Q) \right| \nabla_x \Theta(\omega)) - D_x \Theta(\omega) \left( p \right| \nabla_p T(x, p, Q)) + nT(x, p, Q)
\]

for any \( \omega = (x, p, \zeta, Q, \mathcal{P}, \eta) \in \Omega \). As long as we consider a case in which \( \rho_{\text{Const}}(\mathcal{P}, \zeta, \eta) \) takes the form of \( \rho_2(\mathcal{P}, \zeta) \rho_Y(\eta) \), where \( \rho_2 \) and \( \rho_Y \) are strictly-positive \( C^2 \) functions, such as those that will appear in (S2) (section 4.1), we conclude that \( \text{div} X' = 0 \). Otherwise, in this case, the result \( 0 = \text{div} X'(\omega) = G(x, p, \zeta, Q, \mathcal{P}) + H(\eta) \) for \( \forall \omega \in \Omega \), where \( H(\eta) \equiv mD \ln \rho_Y(\eta) \) becomes the last term of equation (53) and where \( G(x, p, \zeta, Q, \mathcal{P}) \) is the remaining contributions, implies \( \rho_Y(\eta) \propto e^{\text{const} \cdot x} \), which leads to a contradiction to the integrability condition detailed in (C2).

3.2.1. Special types of distributions. Formulae for special types of distributions can be obtained as examples of corollary 7 or theorem 6. The following will be utilized in the numerical simulation (section 5.2).

**Example 8 (marginal distribution).** As for equation (31), since \( A \equiv 1_{B_1 \times \mathbb{R}^n} \) for \( B_1 \in B^D_n \) satisfies (C8), corollary 7 states

\[
\chi_{B_1}^n \circ \pi_x = \chi_{B_1 \times \mathbb{R}^n}^n \circ \pi_R = \left\{ \chi_{B_1 \times \mathbb{R}^n}^n \right\}_R
\]

\[
= \int_{B_1 \times \mathbb{R}^n} \rho_R(x, p) dl_{2n}(x, p)/\int_{D \times \mathbb{R}^n} \rho_R dl_{2n}.
\]

The following two examples are direct consequences of theorem 6.
Example 9 (rewighted distribution). Under condition (C4ii), the current method enables us to produce an arbitrary target density $\rho_{\text{TRG}}(x, p)$, instead of $\rho_k(x, p)$, for the physical system, by a reweighting which needs no explicit form of $\rho_R$. That is,

$$\frac{A\rho_{\text{TRG}}/\rho_{\text{Phys}}}{\langle A \rangle_{\text{TRG}}} = \frac{\int_{D \times \mathbb{R}^n} A\rho_{\text{TRG}} dl_{2n}}{\int_{D \times \mathbb{R}^n} \rho_{\text{TRG}} dl_{2n}} =: \langle A \rangle_{\text{TRG}},$$

(55)

where the functions satisfy (C9) $\rho_{\text{TRG}}: D \times \mathbb{R}^n \to \mathbb{R}^+_1$ is $B_{2n}^{\mathbb{R}^n}$-measurable and $l_{2n}$-integrable, (C10) $A: D \times \mathbb{R}^n \to \mathbb{R}$ is $B_{2n}^{\mathbb{R}^n}$-measurable and $\rho_{\text{TRG}} dl_{2n}$-integrable.

Defining a map $B_R: \Omega \to \mathbb{R}$, $(x, p, \beta) \to (A\rho_{\text{TRG}})(x, p)/\rho_{\text{Phys}}(x, p, \beta)$, we have used an abbreviated form, $A\rho_{\text{TRG}}/\rho_{\text{Phys}}$ in equation (55), to represent the scaled long-time average of $b_R \circ \pi: \Omega \to \mathbb{R}$, $\omega \mapsto (A\rho_{\text{TRG}})(x, p)/\rho_{\text{Phys}}(x, p, \sigma(\omega))$. To show (55), first look at

$$\int_{\Omega_1} [B_R(x, p, \beta)] \rho_{\text{Phys}}(x, p, \beta) dl_{2n+m}(x, p, \beta) \quad = \quad \int_{\Omega_1} [A \rho_{\text{TRG}}](x, p) f(\beta) dl_{2n+m}(x, p, \beta)$$

$$= \int_{D \times \mathbb{R}^n} [A(x, p)] \rho_{\text{TRG}}(x, p) dl_{2n}(x, p) \int_{\sigma(\mathbb{R}^n)} dl_m(\beta) f(\beta) \in \mathbb{R},$$

where we have used (C10) and (C4ii) in the last line to ensure the finiteness. Thus we can apply theorem 6 with $B \equiv B_R$, to obtain

$$\frac{A\rho_{\text{TRG}}/\rho_{\text{Phys}}}{\langle A \rangle_{\text{TRG}}} \equiv [b_R \circ \pi] \in \mathbb{R} \quad (a.e.)$$

(56a)

$$\int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \int_{\sigma(\mathbb{R}^n)} dl_m(\beta) B_R(x, p, \beta) \rho_{\text{Phys}}(x, p, \beta) f(\beta)$$

$$= \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \langle A \rho_{\text{TRG}} \rangle(x, p) \int_{\sigma(\mathbb{R}^n)} dl_m(\beta) f(\beta)$$

$$= \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \langle A \rho_{\text{TRG}} \rangle(x, p) \int_{\sigma(\mathbb{R}^n)} dl_m(\beta) f(\beta)$$

(56b)

$$\int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{R} dl_{2n}$$

$$= \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{R} dl_{2n}$$

(56c)

We can also put $A = 1$ due to (C9), having

$$\rho_{\text{TRG}}/\rho_{\text{Phys}} \equiv [b_R \circ \pi] \in \mathbb{R} \quad (a.e.)$$

(57a)

$$\int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{\text{TRG}}(x, p) \int_{\sigma(\mathbb{R}^n)} dl_m(\beta) f(\beta)$$

$$= \int_{D \times \mathbb{R}^n} dl_{2n}(x, p) \rho_{R} dl_{2n}$$

(57b)

Dividing equation (56) by equation (57) (note the latter is strictly positive) yields equation (55).

Example 10 (distribution regarding $\beta$). For a Borel measurable function $h: \sigma(\mathbb{R}^n) \to \mathbb{R}$, $\beta \mapsto h(\beta)$, such that $\Omega_{1} \to \mathbb{R}$, $(x, p, \beta) \mapsto h(\beta)/\rho_{\text{Phys}}(x, p, \beta) f(\beta)$, is $l_{2n+m}$-integrable, or equivalently, $hfZ$ is integrable, we can apply theorem 6 with putting $B \equiv B_{h}: \Omega_{1} \to \mathbb{R}$, $(x, p, \beta) \mapsto h(\beta)$. Noting $b_{h} \circ \pi(\omega) = B_{h}(x, p, \sigma(\omega)) = h \circ \sigma \circ \pi_{Q}(\omega)$, we thus have
For instance, the distribution of \( Q \) is given by
\[
\mu: B(B_\infty) \rightarrow \mathbb{R}, \quad B \mapsto \chi_{B} \circ \sigma \circ \pi_Q = \frac{\int_{\sigma(\mathbb{R}^n)} f^{\beta}(\beta)Z(\beta)d\mu(\beta)}{\int_{\sigma(\mathbb{R}^n)} f(\beta)Z(\beta)d\mu(\beta)},
\]
which is not just the marginal distribution with respect to \( Q \) represented by \( \mu': B_{m} \rightarrow \mathbb{R}, B' \mapsto \chi_{B'} \circ \pi_{Q'} = P(D \times \mathbb{R}^n \times \mathbb{R} \times B' \times \mathbb{R}^m \times \mathbb{R}) \), but is an induced measure for \( \mu' \) by a map \( \sigma \), that is, \( \mu = \mu' \sigma^{-1} \).

### 4. Example of function setting

For realizing the joint distribution of the physical and parameter systems in a simulation, we should set the specific forms of the functions: \( \rho_{\text{phys}} \) and \( \rho_{\text{Ctrl}} \) in equation (33), \( \sigma \) and \( f \) in equations (33) and (34), and \( T \) in equation (15). These functions should be compatible with (C1)-(C6). First, in section 4.1 we shall consider \( \rho_{\text{phys}}, \rho_{\text{Ctrl}}, \) and \( T \). For \( \sigma \) and \( f \) accordingly, their definitions affect only themselves as described below, and they will be discussed in section 4.2.

#### 4.1. Fundamental functions

We consider \( \rho_{\text{phys}}, \rho_{\text{Ctrl}}, \) and \( T \), while just assuming (C1) and (C4i) for \( \sigma \) and \( f \), respectively. Here we concentrate on an important case such that \( \rho_{\text{phys}}(x, p, \beta) = \rho_{E}(E(x, p), \beta) \), where \( E(x, p) \) is the energy of the physical system. Namely, \( \rho_{\text{phys}} \) depends on \( x \) and \( p \) only through the energy function \( E \). This indicates the situation where the physical state density is determined by the physical-system energy \( E(x, p) \) and the parameter \( \beta \), for which they are independently given. We consider that this corresponds to the situations (I) and (II) introduced in section 1. While for situation (III), where any other system is considered as well as the original physical system, the specification in the form of \( \rho_{\text{phys}}(x, p, \beta) = \rho_{E}(E(x, p), \beta) \) seems to be too restrictive.

The specific details of the functions are given as follows:

\[(S1) \rho_{\text{phys}}: D \times \mathbb{R}^n \times \sigma(\mathbb{R}^m) \rightarrow \mathbb{R}, \quad (x, p, \beta) \mapsto \rho_{E}(E(x, p), \beta), \]

where
\[
E: D \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x, p) \mapsto U(x) + K(p)
\]
with
\[
U: D \rightarrow \mathbb{R}
\]
is of class \( C^2 \).
\( K: \mathbb{R}^n \rightarrow \mathbb{R}, \ p \mapsto (p \mid M^{-1}p)/2, \) with \( M \) being a symmetric, positive-definite square matrix of size \( n \) (over \( \mathbb{R} \));
\[ \rho_E: \mathbb{R}^m \supset D \times \sigma(\mathbb{R}^m) \rightarrow \mathbb{R}, (\epsilon, \beta) \mapsto \rho_E(\epsilon, \beta), \] is of class \( C^3 \), satisfying
\[ \rho_E > 0 \text{ and } D_1\rho_E < 0, \quad \text{(1)} \]
with \( O \) being an open set containing \( E(D \times \mathbb{R}^n) \).

(S2) \( \rho_{\text{Carr}}: \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (P, \zeta, \eta) \mapsto \exp[-K_T(P)]\rho_Z(\zeta)\rho_Y(\eta), \)

where
\[ K_T: \mathbb{R}^m \rightarrow \mathbb{R}, \ P \mapsto (P \mid M_T^{-1}P)/2, \] with \( M_T \) being a symmetric, positive-definite square matrix of size \( m \) (over \( \mathbb{R} \)),
\[ \rho_Z, \rho_Y: \mathbb{R} \rightarrow \mathbb{R}_+, \] which are of class \( C^2 \) and \( L^1 \)-integrable. \(-\text{(2)}\)

(S3) \( T: D \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}, (x, p, Q) \mapsto -\frac{1}{D\ln \rho_E(x, p, \sigma(\mathbb{Q}))}. \)

We see that (C2), (C3), and (C6i) hold. In molecular simulation, potential energy function \( U \) is given as the force field and \( K \) is the kinetic energy with \( M \) such as \( \text{diag}(m_1, \ldots, m_n) \). It follows from (S1) and (S3) that (C6ii) is equivalent to \( L^{2n+m} \)-integrability of a function \( \Omega_T^1 \rightarrow \mathbb{R}, (x, p, \beta) \mapsto f(\beta)D_1\rho_E(E(x, p), \beta), \) which is now a similar condition to (C5), that is, \( (x, p, \beta) \mapsto f(\beta)\rho_E(E(x, p), \beta) \) is \( L^{2n+m} \)-integrable.

Remark 11. Note that \( T(x, p, Q) \) defined by (S3) has the physical dimension of energy or temperature, so that it is consistent with the results stated in section 2.2. Here, we also briefly show how the functional property of \( T(x, p, Q) \) discussed in section 2.2 is realized by adopting (S3). For the first example, we choose the BG density \( \rho_E(\epsilon, \beta) = \exp[-\beta\epsilon] \), considering that \( \rho_E(E(x, p), \sigma(Q)) = \exp[-\sigma(Q)E(x, p)] \), where the dynamical parameter is the inverse heat-bath temperature \( \beta = \sigma(Q) \) (for more details see section 4.2). Then we have \( T(x, p, Q) = 1/\sigma(Q) = 1/\beta \); that is, \( T(x, p, Q) \) is just the heat-bath temperature \( 1/\beta \), where \( \beta \) is not just a constant but a dynamical variable. Namely, the functional property of \( T(x, p, Q) \) is to provide the dynamical heat-bath temperature to the physical system: ‘functional temperature’ \( T(x, p, Q) \) simply becomes the ‘dynamical temperature’ \( 1/\sigma(Q) \) in the BG case. Other possible example is the Tsallis [25] type density \( \rho_E(\epsilon, q) = [1 + (q - 1)\epsilon/T_0]^{1/(1-q)}, \) for which we consider that \( \epsilon = (Q - 1)E(x, p)/T_0 \), where the dynamical parameter we suppose is \( q = \beta \equiv \sigma(Q) \in (0, 1] \) (while \( T_0 \) > 0 is fixed to be a constant parameter defined in the Tsallis distribution). In this case we have \( T(x, p, Q) = T_0 + (\sigma(Q) - 1)E(x, p) \), that is, the functional temperature \( T(x, p, Q) \) provides a temperature-like quantity composed of the constant part \( T_0 \) and the fluctuating part yielded by the dynamical parameter \( q = \sigma(Q) \) along with \( E(x, p) \). It should be noted that insightful ideas regarding a general temperature function are seen in [37, 38].

Now, applying (S1) and (S2) into equation (33) and substituting the resultant density and temperature-function (S3) into equation (15), we have the following EOM:
\[
\begin{align*}
\dot{x} &= M^{-1}p \in \mathbb{R}^n, \\
\dot{\rho} &= -\nabla U(x) - T(x, p, Q) \tau_{\mathbb{Z}}(\zeta) p \in \mathbb{R}^n, \\
\dot{\zeta} &= 2K(p) - nT(x, p, Q) \in \mathbb{R}^1, \\
\dot{\mathcal{Q}} &= M^T\mathcal{P} \in \mathbb{R}^m, \\
\dot{\mathcal{P}} &= -\nabla \mathcal{U}_{E(x, p)}(Q) - \tau_Y(\eta) \mathcal{P} \in \mathbb{R}^m, \\
\dot{\eta} &= 2K_\eta(\mathcal{P}) - m \in \mathbb{R}^1,
\end{align*}
\]

where \(\tau_{\mathbb{Z}}(\zeta) \equiv -D \ln \rho_{\mathbb{Z}}(\zeta), \tau_Y(\eta) \equiv -D \ln \rho_Y(\eta),\) and
\[
\mathcal{U}(Q) \equiv -\ln \left[ \rho_f(Q)\left|D_1\rho_E(\epsilon, \sigma(Q))\right|\right].
\]

Here, \((x, p, \zeta)\) forms the NH equations with the functional temperature \(T(x, p, Q)\), instead of a constant external temperature, along with the temperature-dependent friction \(T(x, p, Q) \tau_{\mathbb{Z}}(\zeta)\). In addition, \((Q, \mathcal{P}, \eta)\) takes again the NH form with parameter-system potential energy, \(\mathcal{U}_{E(x, p)}(Q)\), which is a function of parameter coordinates \(Q\) and depends also on the physical-system energy \(\epsilon \equiv E(x, p)\), along with a friction \(\tau_Y(\eta)\) and a unit temperature. In this respect, we call this coupled NH EOM.

**Remark 12.** Since (S1) contains potential function \(U(x)\), it may be natural to investigate the dependence of the key quantities upon a shift of the origin of \(U\), namely, \(U \rightarrow U + u_0\) for constant \(u_0 \in \mathbb{R}\). Such key quantities are \(\rho\) (equation (33)), that space average \(\langle g \rangle\) of any \(g\) (equation (21)), and the EOM (equation (60)). We can find a simple dependence upon \(u_0\) of the behaviors of the quantities, if we restrict our attention to the case that there exists a function \(\lambda: \mathbb{R} \times \sigma(\mathbb{R}^m) \rightarrow \mathbb{R}_+, \) such that
\[
\rho_E(\epsilon + u_0, \beta) = \lambda(u_0, \beta)\rho_E(\epsilon, \beta)
\]

for all possible \(\epsilon, u_0,\) and \(\beta\) (that is, the ratio, \(\rho_E(\epsilon + u_0, \beta)/\rho_E(\epsilon, \beta)\), is independent of \(\epsilon\)). Such an example of \(\rho_E\) will be seen in (S1-1) below (set \(\lambda(u_0, \beta) = \exp(-\beta u_0)\)). In this case, \(\rho \mid_{U \rightarrow U + u_0}(\omega) = \rho(\omega)\lambda(u_0, \sigma(Q)) \equiv \rho(\omega)\lambda_{u_0}(\sigma(Q))\) for all \(\omega \in \Omega\), or simply,
\[
\rho \mid_{U \rightarrow U + u_0} = \rho \mid_{-\beta\lambda_{u_0}},
\]

that is, the shift for \(U\) in \(\rho\) is equivalent to a factorization of \(f\) (see equations (33) and (34)). Despite the fact that \(T\) in (S3) depends on \(E\), we see the invariance:
\[
T \mid_{U \rightarrow U + u_0} = T.
\]
Thus, the vector field \(X'\) (defining equation (60)), which is made from \(\rho\) and \(T\), obeys
\[
X' \mid_{U \rightarrow U + u_0} = X' \mid_{-\beta\lambda_{u_0}}.
\]

Hence, if we have the converse factorization \(f \rightarrow f/\lambda_{u_0}\) as well as the shift \(U \rightarrow U + u_0\), then we recover an invariance for the density, \(\rho \mid_{U \rightarrow U + u_0} = \rho\), together with the invariances for \(\langle g \rangle\) and \(X'\). Note that these invariances are a matter of theoretical interest. In the applications, the relationships, (63)–(65), are really concerned, since \(\rho\) defines the realized density, \(T\) specifies the long-time averages (equations (22) and (29a)), and \(X'\) defines the EOM (equation (60)).

**4.2. Remaining functions**

We here consider subordinate functions defined in (S1) and (S2), the transform function \(\sigma\), and \(\beta\)'s density function \(f\). As a natural choice we can set these functions as follows:

---

4.2. Remaining functions

We here consider subordinate functions defined in (S1) and (S2), the transform function \(\sigma\), and \(\beta\)'s density function \(f\). As a natural choice we can set these functions as follows:

---

I Fukuda and K Moritsugu

J. Phys. A: Math. Theor. 48 (2015) 455001
(S1-1) \( \rho_E: \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}^+ \), \( (\epsilon, \beta) \mapsto \exp[-\beta \epsilon] \), with \( m = 1 \),
(S2-1) \( \rho_z: \mathbb{R} \rightarrow \mathbb{R}^+ \), \( \zeta \mapsto \exp[-c_z \zeta^2] \), with \( c_z > 0 \),
(S2-2) \( \rho_y: \mathbb{R} \rightarrow \mathbb{R}^+ \), \( \eta \mapsto \exp[-c_y \eta^2] \), with \( c_y > 0 \),
(S4) \( m = 1 \), and \( \sigma: \mathbb{R} \rightarrow \mathbb{R}^+ \), \( Q \mapsto c \exp[Q/l] \), with \( c, l > 0 \).

Here \( c_z, c_y, c, \) and \( l \) are real parameters. \( \rho_E \) is a usual Boltzmann–Gibbs density, and \( \rho_z \) and \( \rho_y \) are Gaussians, which are familiar forms having good properties. In fact, they are \( C^\infty \), and both (\( \bullet \) 1) (in (S1)) and (\( \bullet \) 2) (in (S2)) are satisfied. \( \sigma \) in (S4) ensures \( \beta \equiv \sigma(Q) > 0 \) for any \( Q \in \mathbb{R} \).

We also see the validity of (C1) (where \( \det D\sigma > 0 \)), and thus (C6i) holds. In fact, (S3) obeying (S1-1) signifies
\[
T(x, p, Q) = 1/\sigma(Q). \tag{66}
\]

As stated in remark 11, the physical meaning of equation (66) is clear: the functional temperature \( T(x, p, Q) \) becomes the dynamical (heat-bath) temperature \( 1/\sigma(Q) \). Thus we are considering situation (I): temperature \( 1/\beta = 1/\sigma(Q) \) is fluctuated obeying a certain distribution \( f \).

The remaining setting is for \( f \), and we set \( f \equiv f_G \), defined by
\[
(S5) \ f_G: \mathbb{R}^+ \rightarrow \mathbb{R} \), \( \beta \mapsto \frac{\beta^\alpha}{\Gamma(\alpha)} \exp[-d/\beta] \) with \( \alpha, d > 0 \), where \( \Gamma \) is the gamma function.

Then we see that (C4) is satisfied ((ii) holds as \( \int_{\mathbb{R}^+} f_G(\beta) d\beta = 1 \)). The remaining conditions, (C5) and (C6ii), are not trivial to confirm their validities in general. Note that, as is shown by lemma 5 with putting \( b \equiv 1/T \), (C6ii) can be replaced with the condition that \( \langle 1/T \rangle < \infty \), that is, the inverse temperature function has a finite average. Considering this fact, (C6ii) seems to be a physically natural requirement, even if it may be unnecessary under a certain alternative consideration (e.g., considering a transformation of the simple long-time average instead of considering the transformed measure \( P' \)). For a toy model system, the validities of (C5) and (C6ii) are confirmed, and in fact they are valid for a 1D harmonic oscillator model used in the current numerical study (see section 5.2.1) as long as we set
\[
\alpha > 1. \tag{67}
\]

The choice of (S5) for \( f \) here is due to both these reasons and physically highly interesting features coming from the intimate relationship with nonextensive statistical mechanics [17, 21, 22]. For general systems, we assume the validities of (C5) and (C6ii).

5. Numerics

5.1. Numerical integrator

We introduce a numerical integration scheme for the ODE, (60). To do this, we define an extended ODE, according to the scheme in [34, 39], on an extended space \( \Omega' \equiv \Omega \times \mathbb{R}^+ \):
\[
\begin{alignat}{2}
\dot{\omega} &= X'(\omega), \\
\dot{v} &= Y(\omega),
\end{alignat} \tag{68a}
\]
for \( \omega \in \Omega \). Here, equation (68a) represents equation (60), \( v \) is introduced as an extended variable on \( \mathbb{R} \), and
\[ \mathcal{J}(\omega) \equiv -T(x, p, Q) \nabla X(\omega), \]

with \( X(\omega) = X'(\omega)/T(x, p, Q) \) \( (\forall \omega \in \Omega) \) in equation (69). Then, ODE (68) has an invariant function [34] defined by

\[ L: \Omega' \to \mathbb{R}, \omega' \equiv (\omega, v) \mapsto -\ln \rho(\omega) + v. \]

Thus the numerical error of the integration can be checked by monitoring the value of equation (70) in the integration process.

For ODE (68) we can use a natural decomposition of its field, \( \Omega' \to \mathbb{R}^{N+1}; \omega' \mapsto (X'(\omega), \mathcal{J}(\omega)) \), by assuming that \( T(x, p, Q) \) is independent of \( p \), which is satisfied e.g. in the (S1-1) case (see equation (66)). Namely we take account of the forms,

\begin{align*}
S_1(x, Q, \zeta) &\equiv T(x, p, Q) \tau_2(\zeta), \quad (71) \\
S_2(x, Q) &\equiv nT(x, p, Q), \quad (72)
\end{align*}

for the 2nd and 3rd equations for equation (60). Then the fundamental maps corresponding to individual fields are easily obtained [34], and thus a symmetric, desired order numerical integrator can be constructed. The fundamental maps are

\begin{align*}
\Phi_1^{(1)}(\omega') &= (tM^{-1}p + x, p, \zeta, Q, P, \eta, v), \quad (73a) \\
\Phi_1^{(2)}(\omega') &= (x, tF(x) + p, \zeta, Q, P, \eta, v), \quad (73b) \\
\Phi_1^{(3)}(\omega') &= (x, p, [\kappa(p) - S_2(x, Q)]t + \zeta, Q, P, \eta, v), \quad (73c) \\
\Phi_1^{(4)}(\omega') &= (x, \exp[-tS_1(x, Q)]p, \zeta, Q, P, \eta, v), \quad (73d) \\
\Phi_1^{(71)}(\omega') &= (x, p, \zeta, tM^{-1}P + Q, P, \eta, v), \quad (73e) \\
\Phi_1^{(72)}(\omega') &= (x, p, \zeta, Q, tF(x, p, Q) + P, \eta, v), \quad (73f) \\
\Phi_1^{(73)}(\omega') &= (x, p, \zeta, Q, \eta, t[2K_T(P) - m] + \eta, v), \quad (73g) \\
\Phi_1^{(74)}(\omega') &= (x, p, \zeta, Q, \exp[-tR_N(\eta)]P, \eta, v), \quad (73h) \\
\Phi_1^{(5)}(\omega') &= (x, p, \zeta, Q, \eta, t\mathcal{J}(\omega) + v), \quad (73i)
\end{align*}

where \( F(x) \equiv -\nabla U(x) \) and \( F_T(x, p, Q) \equiv -\nabla U_{E(x,p)}(Q) \) are the forces for the physical system and the temperature system, respectively. The simplest symmetric, second-order integrator with time step \( h \) is given, e.g., by

\begin{align*}
\Phi_1^{(5)} \circ \Phi_1^{(4)} \circ \Phi_1^{(3)} \circ \Phi_1^{(2)} \circ \Phi_1^{(71)} \circ \Phi_1^{(72)} \circ \Phi_1^{(73)} \circ \Phi_1^{(4)} \circ \Phi_1^{(3)} \circ \Phi_1^{(2)} \circ \Phi_1^{(1)} \\
\circ \Phi_1^{(2)} \circ \Phi_1^{(3)} \circ \Phi_1^{(4)} \circ \Phi_1^{(3)} \circ \Phi_1^{(2)} \circ \Phi_1^{(1)}.
\end{align*}

Here, although \( \Phi_1^{(72)} \) requires the value of \( U(x) \) in general through the evaluation of \( F_T(x, p, Q) \) (and \( \Phi_1^{(3)} \) and \( \Phi_1^{(4)} \) may also require it through the evaluation of \( T(x, p, Q) \) in a general case), the above map ordering ensures that the evaluations of the potential energy \( U(x) \) and the force \( F(x) \) is once in this integrator mapping. Maps, \( \Phi_1^{(5)} \), at both ends do not need this kind of evaluation but just refer to the values. Note that \( S_1 \) and \( S_2 \) can be used for an EOM that has a more generalized form than equation (60)); and equation (73) can be extended to apply the scheme into such a generalized form.
5.2. Numerical test

To verify the present method, we performed a numerical simulation using the function setting described in section 4 with

\[ E(x, p) = \frac{1}{2} x^2 + \frac{1}{2} p^2 \]

for \((x, p) \in \mathbb{R}^2\) (that is, \(n = 1\)). This provides the simplest model system, where a one-dimensional harmonic oscillator (1HO) is in a heat bath whose temperature is fluctuating obeying the distribution density \(f \equiv f_G\) (namely, type (I), where the physical system is defined by the 1HO and the parameter system is the temperature system). Note that the 1HO is suitable to validate the method because it gives a typical and important model that describes the physical system behavior around the equilibrium state and because it is simple enough such that the exact distribution is analytically obtained. There also includes a complexity in that the BG distribution is not trivially achieved via the sampling by use of the single NH equations due to the lack of ergodicity [40].

Our main purpose of the simulation is to examine the accuracy of the generated distributions of both the physical and temperature variables. Since this method presents new dynamics including both the physical and temperature variables, we also observed the associated trajectories generated by the simulation. The correlations among the dynamical temperature, the kinetic energy and the potential energy were also investigated.

5.2.1. Protocol. For numerical simulation, we used the functions defined in (S1)–(S5). The following parameter values were used: \(M_F = 1\) (see (S2)), \(c_x = c_y = 1\) (see (S2-1) and (S2-2)), and \(c = 1\) (see (S4)). For (S5), we put \(a = d = 4\), so that equation (67) is satisfied. We used \(l = 2.24\) to set \(1 \sim (\alpha + 1)/l^2\) (see (S4)). We integrated \(10^8\) time steps using equation (74) with a unit time of \(1 \times 10^{-3}\).

The distribution of each variable was obtained by the corresponding marginal distribution. For example, the distribution of \(x\) is represented by equation (32) with \(B_t\) being each bin, where the theoretical distribution is represented by the rhs of the equation, which turns out to be the rhs of equation (54), and the simulated distribution was estimated by the lhs, \(\frac{1}{B_t} \sum \rho_{\text{TRG}}\), with a finite-time approximation. Similar procedures were taken for \(p, \zeta, \mathcal{P}, \) and \(\eta\). The distribution of \(\beta \equiv \sigma(Q)\) was obtained from equation (59) for each bin \(B\). The reweighting to the BG distribution density, \(\rho_{BG} : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, p) \mapsto e^{-\beta_{BG} E(x, p)}\), at \(\beta_{BG} = 1\) was done by equation (55), where (C9) is obviously satisfied with \(\rho_{TRG} \equiv \rho_{BG}\) for any \(\beta_{BG} > 0\) and \(A \equiv \chi^2_{BG}\) satisfies (C10) for any bin \(B \in B_2\).

5.2.2. Results. We show numerical results to confirm the realization of the distribution, equation (33) with (S1)–(S5), via equation (60). Figure 1 shows the distribution densities of the coordinate \(x\), the momentum \(p\), and the control variable \(\zeta\) for the physically-based system, and the inverse dynamical temperature \(\beta = \sigma(Q)\), the momentum \(\mathcal{P}\), and the control variable \(\eta\) for the temperature system. The simulated and exact results agree well, as was seen in small standard deviations of the discrepancies from the exact distributions for \(x, p, \zeta, \beta, \mathcal{P}, \) and \(\eta\), being \(2.3 \times 10^{-3}\), \(2.2 \times 10^{-3}\), \(3.6 \times 10^{-3}\), \(4.5 \times 10^{-3}\), \(2.4 \times 10^{-3}\), and \(3.2 \times 10^{-3}\), respectively. These results indicate that the present method produced a sufficiently accurate distribution. The reweighted result to the BG distributions was also satisfactory, where the standard deviations of the discrepancies from the exact distributions for \(x\) and \(p\) were \(3.1 \times 10^{-3}\) and \(3.0 \times 10^{-3}\), respectively.
The behavior of the physical system in the nonequilibrium environment provided by the temperature system is shown in Figure 2(a) and (b). The time developments of physical temperature, $T_P \equiv 2K(p)/n = p^2$, are perturbed from the pure 1HO Hamiltonian system, which should have period $\pi$ for $p^2$. Compared with $T_P$, the dynamical temperature, $T_D \equiv \sigma(Q)^{-1} = T(x, p, Q)$, changes more slowly (figure 2(a)). Around the time when $T_D$ is large in oscillating-like motions, the amplitudes of $T_P$ also become larger in an interesting manner. As shown in figures 2(d)–(f), $\zeta$, $\varphi$, and $\eta$ show oscillation-like behaviors in a nonlinear regime. As expected from the EOM, $\dot{\zeta} = T_P - T_D$ (the 3rd equation in (60)), figure 2(d) shows that $\zeta$ mainly obeys $T_P$, which shows faster oscillating motion than that of $T_D$. On the other hand, $\varphi$ mainly changes according to $T_D$ (figure 2(e)), as suggested by the relations, $\varphi = Q$ and $T_D \propto -Q/|Q|$. Similar correspondence appears in $\eta$ (figure 2(f)) and $\varphi$, via $\eta = \beta^2 - 1$. Figure 2(b) shows a time development of physical potential energy $U(x)$, suggesting a correlated motion with $T_P$. In fact, the correlation with a retarded time is seen by a correlation function (figure 2(c)), defined by $G(t) \equiv \sum_{t_0} K(p(t_0))U(x(t + t_0))$. These correlations must be due to the nonlinear feature of ODE (60) and help to surmount nonergodic behaviors seen in conventional methods [41]. Keeping constant in the trajectory of the invariant function (equation (70)) in figure 2(b) shows the success in the numerical integration.

6. Conclusion

We have developed a new formalism, DDD, to realize the densities of both the physical and parameter systems through the scaled long-time average generated by an ODE. The joint distribution density of the physical system and the parameter system is realized under the ergodic assumption of the flow, and the physical system density $\rho_R$ is realized without its
Figure 2. Time development of quantities: (a) physical temperature $T_p = p^2$ and dynamical temperature $T_d = \sigma(Q)^{-1}$; (b) potential energy $U$ and invariant $L$; (c) time-correlation between $U$ and $T_p$; (d) control variable $\zeta$ for the physical system; (e) momentum $P$ and (f) control variable $\eta$ for the temperature system.

explicit form (that is, the superstatistics distribution is dynamically realized). The functions employed here to specify the vector field are simply an example among many choices. DDD demonstrates its potential for being (i) fruitfully utilized for various problems via the creation
of arbitrarily designed temperature functions ((S3)), (ii) exerted for well-planned parameters instead of temperature (equations (15) and (60)), and (iii) generalized for addressing multiple densities hierarchically for, e.g. multiscale problems (equation (12)). These new, general ideas are utilized for efficiently simulating a physical system in nonequilibrium and in various environments, and thus it aims to advance molecular simulation, which is now a standard tool for atomic, molecular, soft-matter, and biological physics. We have focused on a theoretical study, and extensive applications are now in progress.

Acknowledgments

I F is grateful for a Grant-in-Aid for Scientific Research (C) (25390156) and K M was supported by a Grant-in-Aid for Young Scientists (15K18520), both from JSPS. I F was partially supported by the ‘Development of core technologies for innovative drug development based upon IT’ from Japan Agency for Medical Research and development, AMED.

References

[1] Allen M and Tildesley D 2002 *Computer Simulation of Liquids* (New York, Oxford)
[2] Hoover W G 1991 *Computational Statistical Mechanics* (Amsterdam: Elsevier)
[3] Evans D J and Morriss G 1990 *Statistical Mechanics of Nonequilibrium Liquids* (London: Academic Press)
[4] Mitsutake A, Sugita Y and Okamoto Y 2001 Generalized-ensemble algorithms for molecular simulations of biopolymers *Biopolymers* **60** 96–123
[5] Nadler W and Hansmann U 2007 Generalized ensemble and tempering simulations: a unified view *Phys. Rev. E* **75** 026109
[6] Cho K and Joannopoulos J 1992 Ergodicity and dynamical properties of constant-temperature molecular dynamics *Phys. Rev. A* **45** 7089–97
[7] di Tolla F D and Ronchetti M 1993 Applicability of Nosé isothermal reversible dynamics *Phys. Rev. E* **48** 1726–37
[8] Fukuda I and Nakamura H 2002 Tsallis dynamics using the Nosé-Hoover approach *Phys. Rev. E* **65** 026105
[9] Fukuda I, Horie M and Nakamura H 2005 Deterministic design for Tsallis distribution sampling *Chem. Phys. Lett.* **405** 364–70
[10] Moritsugu K, Terada T and Kidera A 2010 Scalable free energy calculation of proteins via multiscale essential sampling *J. Chem. Phys.* **133** 224105
[11] Hoover W G, Ladd A J C and Moran B 1982 High-Strain-Rate Plastic Flow Studied via Nonequilibrium Molecular Dynamics *Phys. Rev. Lett.* **48** 1818–20
[12] Evans D J, Hoover W G, Failor B H, Moran B and Ladd A J C 1983 Nonequilibrium molecular dynamics via Gauss’s principle of least constraint *Phys. Rev. A* **28** 1016–21
[13] Nosé S 1984 A unified formulation of the constant temperature molecular dynamics methods *J. Chem. Phys.* **81** 511–9
[14] Hoover W G 1985 Canonical dynamics: equilibrium phase-space distributions *Phys. Rev. A* **31** 1695–7
[15] Hünenberger P H 2005 Thermostat algorithms for molecular dynamics simulations *Advanced Computer Simulation: Approaches for Soft Matter Sciences I, Advances in Polymer Science* vol 173 ed C Holm and K Kremer (Berlin: Springer) 105–49
[16] Jepps O G and Rondoni L 2010 Deterministic thermostats, theories of nonequilibrium systems and parallels with the ergodic condition *J. Phys. A: Math. Theor.* **43** 133001
[17] Beck C 2001 Dynamical Foundations of Nonextensive Statistical Mechanics *Phys. Rev. Lett.* **87** 180601
[18] Beck C and Cohen E G D 2003 Superstatistics *Physica A* **322** 267–75
[19] Hanel R, Thurner S and Gell-Mann M 2011 Generalized entropies and the transformation group of superstatistics *Proc. Natl. Acad. Sci.* **108** 6390–4

[20] Abe S, Beck C and Cohen E 2007 Superstatistics, thermodynamics, and fluctuations Phys. Rev. E 76 031102
[21] Wilk G and Wlodarczyk Z 2000 Interpretation of the nonextensivity parameter q in some
applications of Tsallis statistics and Lévy distributions Phys. Rev. Lett. 84 2770–3
[22] Tsallis C 1988 Possible generalization of Boltzmann-Gibbs statistics J. Stat. Phys. 52 479–87
[23] Tsallis C, Mendes R and Plastino A R 1998 The role of constraints within generalized
nonextensive statistics Physica A 261 534–54
[24] Abe S and Rajagopal A K 2003 Revisiting disorder and Tsallis statistics Science 300 249–50
[25] Tsallis C 2009 Introduction to Nonextensive Statistical Mechanics (Heidelberg: Springer)
[26] Reynolds A 2003 Superstatistical mechanics of tracer-particle motions in turbulence Phys. Rev. Lett. 91 084503
[27] Abe S and Thurner S 2005 Complex networks emerging from fluctuating random graphs: analytic
formula for the hidden variable distribution Phys. Rev. E 72 036102
[28] Baiesi M, Paczuski M and Stella A 2006 Intensity thresholds and the statistics of the temporal
occurrence of solar flares Phys. Rev. Lett. 96 051103
[29] Beck C 2009 Superstatistics in high-energy physics Eur. Phys. J. A 40 267–73
[30] Abul-Magd A Y, Akemann G and Vivo P 2009 Superstatistical generalizations of Wishart-
Laguerre ensembles of random matrices J. Phys. A: Math. Theor. 42 175207
[31] García-Morales V and Krischer K 2011 Superstatistics in nanoscale electrochemical systems Proc. Natl. Acad. Sci 108 19535–9
[32] Gao Y Q 2008 An integrate-over-temperature approach for enhanced sampling J. Chem. Phys. 128 064105
[33] Hahn M G, Jiang X and Umarov S 2010 On, q-Gaussians and exchangeability J. Phys. A: Math. Theor 43 165208
[34] Fukuda I and Nakamura H 2006 Construction of an extended invariant for an arbitrary ordinary
differential equation with its development in a numerical integration algorithm Phys. Rev. E 73 026703
[35] Fukuda I 2010 Comment on ‘Preserving the Boltzmann ensemble in replica-exchange molecular
dynamics’ J. Chem. Phys. 132 127101
Fukuda I 2010 J. Chem. Phys. 132 127101
[36] Fukuda I and Nakamura H 2004 Efficiency in the generation of the Boltzmann-Gibbs distribution
by the Tsallis dynamics reweighting method J. Phys. Chem. B 108 4162–70
Fukuda I and Nakamura H 2005 Molecular dynamics sampling scheme realizing multiple
distributions Phys. Rev. E 71 046708
[37] Jepps O G, Ayton G and Evans D J 2000 Microscopic expressions for the thermodynamic
temperature Phys. Rev. E 62 4757–63
[38] Samoletov A A, Dettmann C P and Chaplain M A J 2007 Thermostats for Slow Configurational
Modes J. Stat. Phys. 128 1321–36
[39] Queyroy S, Nakamura H and Fukuda I 2009 Numerical examination of the extended phase-space
volume-preserving integrator by the Nosé-Hoover molecular dynamics equations J. Comput. Chem. 30 1799–815
[40] Posch H, Hoover W and Vesely F 1986 Canonical dynamics of the Nosé-oscillator: Stability,
order, and chaos Phys. Rev. A 33 4253–65
[41] Legoll F, Luskin M and Moeckel R 2006 Non-ergodicity of the Nosé-Hoover thermostatted
harmonic oscillator Arch. Rational Mech. Anal. 184 449–63