Ann. Phys. (Berlin) **528**, No. 5, 381–393 (2016) / DOI 10.1002/andp.201500221

annalen der physik

Original Paper

Statistical mechanics of holonomic systems as a Brownian motion on smooth manifolds

Fabio Manca[1], *Pierre-Michel Déjardin*[2], *and Stefano Giordano*[1,3,*]

Received 23 June 2015, revised 9 October 2015, accepted 20 October 2015
Published online 9 November 2015

The statistical mechanics of arbitrary holonomic scleronomous systems subjected to arbitrary external forces is described by specializing the Lagrange and Hamilton equations of motion to those of the Brownian motion on a manifold. In this context, the Klein-Kramers and Smoluchowski equations are derived in covariant form, and it is demonstrated that these equations have equilibrium solutions corresponding to the Gibbs distribution, in agreement with standard thermodynamics. At last, the Langevin dynamics corresponding to the Smoluchowski limit is found to exactly correspond to the Brownian motion on a smooth manifold. These results find significant applications in the study of several statistical properties of constrained molecular assemblies (e.g. polymers) of interest in chemistry, physics and biology.

1 Introduction

The observation of the motion of small pollen particles immersed in water [1, 2], and the analysis of the hydrodiffusion through membranes [3], played a crucial role for the development of the statistical physics at the beginning of the nineteenth century. The modern kinetic molecular interpretation of both phenomena was confirmed through the pioneering theoretical analysis of diffusion [4, 5], the introduction of stochastic processes [6] and the experimental test of the atomic hypothesis [7, 8].

In the early twentieth century the Brownian motion was described directly through the Newtonian dynamic equations, suitably modified by two additional force terms [9]: a dissipative friction force and a random contribution representing the fluctuating effects of the collisions. Then, the resulting Langevin stochastic differential equation was studied analysing the time evolution of the density probability of the system state through the so-called Fokker-Planck equation [10–14]. Typically, this equation written within the entire phase space (coordinates and velocities) of a mechanical system is referred to as the Klein-Kramers equation. The analysis of several linear and nonlinear Brownian or stochastic systems can be found in literature [15–21]. Moreover, the ideas underlying these approaches have been also applied to quantum mechanics, where the density probability is substituted by the reduced density matrix [22–25].

Many important applications of the Brownian motion theory can be found in soft matter [26] and polymer physics [27, 28]. For instance, the thermo-elastic properties of DNA [29, 30] and other polymer chains [31–34] have been largely investigated. Also, models for polymers undergoing conformational transitions have been developed [35–38]. Further, we can cite the biological motion at the molecular scale generated by molecular motors, constituted by a periodic asymmetrical potential combined with a ratchet effect [39, 40]. Another important application is given by the dynamics of the magnetization vector in ferromagnetic particles. The classical Landau-Lifshitz-Gilbert equation [41, 42], generalized by adding a noise term representing the magnetic fluctuations [43–46], leads to stochastic models of the magnetization dynamics [21, 47–49].

All provided examples have in common that the systems involved in such dynamics are subjected to holonomic constraints, and sometimes their motion cannot be described by orthogonal curvilinear coordinates. Consequently, these systems must be analysed through the general framework of analytical mechanics invoking the introduction of generalized coordinates. It

* Corresponding author E-mail: Stefano.Giordano@iemn.univ-lille1.fr
1 Institute of Electronics, Microelectronics and Nanotechnology (UMR CNRS 8520), F-59652 Villeneuve d'Ascq, France
2 Laboratoire de Mathématiques et Physique (EA 4217), Université de Perpignan Via Domitia, F-66860, Perpignan, France
3 International Associated Laboratory LEMAC/LICS, ECLille, F-59652 Villeneuve d'Ascq, France

is important to remark that, to introduce the statistical behavior of these systems, we have to define the collision mechanism governing the non-equilibrium dynamics and the relaxation processes. In this work, we assume the Langevin scheme, which is a simple idealization of the actual situation leading to efficient models [21]. Coherently, we consider the standard analytical mechanics, rewritten by taking into account two additional terms: a friction and a noise term. They describe the interaction between the system and a thermal bath, introducing an energy exchange and a proper asymptotic thermodynamic equilibrium (Gibbs distribution).

The objective of this work is to develop a Klein-Kramers equation for a general holonomic scleronomous system, to prove the coherence with the asymptotic Gibbs distribution and to obtain the overdamped limit yielding the so-called Smoluchowski equation. Both the generally covariant versions of the Klein-Kramers and the Smoluchowski equations are presented, proving that non-equilibrium statistical mechanics can be interpreted with a Brownian motion on a Riemannian manifold. For a system with N particles, following Langevin, we have introduced $3N$ noise terms to describe the fluctuations. However, the constrained system is represented by $n < 3N$ degrees of freedom. By means of the reduction of the noise terms from $3N$ to n, the final dynamics is naturally described on the manifold (dimension n) without the necessity for an embedding in a larger space (dimension $3N$). This is achieved through the analysis of the statistical equivalence of a class of specific Langevin equations. Finally, for the overdamped case (high friction limit) we obtain an equation of motion, which exactly corresponds to the equation postulated within the mathematical community for the description of Brownian motion on a manifold.

2 Lagrangian approach

We consider a system composed of N particles with masses $m_i (i = 1, \ldots, N)$ subjected to holonomic scleronomous constraints. This means that such constraints concern only the positions of the particles and they not depend on time [50]. If the vectors $\vec{r}_1, \ldots, \vec{r}_N$ identify the cartesian coordinates of the particles, the constraints can be written as $f_\alpha(\vec{r}_1, \ldots, \vec{r}_N) = 0$ ($\forall \alpha = 1, \ldots, p$). So, the degrees of freedom of the system are $n = 3N - p$. We define $(\vec{r}_1, \ldots, \vec{r}_N) = (\xi_1, \ldots, \xi_{3N})$ and, without loss of generality, we consider $\xi \triangleq (\xi_{p+1}, \ldots, \xi_{3N})$ as the vector of free variables. Hence, we introduce the generalized coordinates $q \triangleq (q^1, \ldots, q^n)$ through an arbitrary transformation $\xi \leftrightarrow q$, with $\det(\partial \xi / \partial q) \neq 0$. It means that any particle position can be described by $\vec{r}_i = \vec{r}_i(q^1, \ldots, q^n)$ and the corresponding velocities can be determined as follows

$$\vec{v}_i = \frac{d}{dt}\vec{r}_i = \frac{\partial \vec{r}_i}{\partial q^k}\dot{q}^k. \tag{1}$$

We assume the implicit Einstein convention for the sums over the indices running on the n degrees of freedom. However, we maintain explicit notation for the sums over the N particles. The Newton motion equations are therefore

$$\vec{F}_i + \vec{\Phi}_i = m_i \vec{a}_i \qquad (\forall i = 1, \ldots, N), \tag{2}$$

where $\vec{F}_i$ are the external forces applied to the system, $\vec{\Phi}_i$ are the reaction forces (applied by the constraints to the system), and $\vec{a}_i$ are the acceleration vectors. From the virtual work principle $\sum_{i=1}^{N} \vec{\Phi}_i \cdot \vec{v}_i = 0$ [50], combined with Eqs. (1) and (2), we obtain

$$\sum_{i=1}^{N} \left(m_i \vec{a}_i - \vec{F}_i \right) \cdot \frac{\partial \vec{r}_i}{\partial q^k}\dot{q}^k = 0, \tag{3}$$

where the sum over k must be highlighted. Indeed, since the q^k's are independent and arbitrary, we have that

$$\sum_{i=1}^{N} \left(m_i \vec{a}_i - \vec{F}_i \right) \cdot \frac{\partial \vec{r}_i}{\partial q^k} = 0 \quad (\forall k = 1, \ldots, n). \tag{4}$$

These steps, very well known in standard analytical mechanics (see, e.g., [50]), are reported here only to introduce all the relevant quantities. By defining the kinetic energy

$$T = \sum_{i=1}^{N} \frac{1}{2} m_i \vec{v}_i \cdot \vec{v}_i = \frac{1}{2} a_{kh}(q) \dot{q}^k \dot{q}^h, \tag{5}$$

where

$$a_{kh}(q) = \sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q^k} \cdot \frac{\partial \vec{r}_i}{\partial q^h}, \tag{6}$$

the classical Lagrangian approach proves that Eq. (4) is equivalent to

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^k} \right) - \frac{\partial T}{\partial q^k} = \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q^k} \triangleq Q_k, \tag{7}$$

where $Q_1, \ldots, Q_n$ are the so-called generalized forces [50]. The tensor quantity defined in Eq. (6) is very important since it will assume the role of metric tensor for a given Riemannian manifold, perfectly representing the geometry of the system. Now, we suppose that $\vec{F}_i$ takes into account the following terms: (i) a conservative force field applied to the system, (ii) a friction force mimicking the energy transfer from the particles to the thermal bath, and (iii) a noise term mimicking the energy transfer from the bath to the system. We then postulate

$$\vec{F}_i = -m_i \beta \vec{v}_i + \sqrt{D m_i}\,\vec{n}_i(t) - \frac{\partial V}{\partial \vec{r}_i}, \tag{8}$$

where β is the friction coefficient (per unit mass) and D is the diffusion coefficient (per unit mass). As usual, we assume the following hypotheses on the noises: $\vec{n}_i(t) \in \mathbb{R}^3$ are Gaussian stochastic processes, $E\{\vec{n}_i(t)\} = 0$, and $E\{\vec{n}_i(t_1) \otimes \vec{n}_j^T(t_2)\} = 2\delta_{ij}\mathbf{I}_3\delta(t_1 - t_2)$ (here E means "expected value", T means "transposed", δ_{ij} is the Kronecker delta, $\delta(.)$ is the Dirac delta function, $\otimes$ is the tensor product of vectors and $\mathbf{I}_3$ is the 3×3 identity matrix). We will prove that these properties are sufficient to obtain a correct thermodynamic behavior for the system. The stochastic differential equation (SDE) obtained in Eqs. (7) and (8) merits a further discussion: from the mathematical point of view there are several different approaches for defining the meaning of a SDE, such as the Itô stochastic calculus, the Stratonovich one or the backward (or anti-Itô) formulation [51, 52]. In this regards, one should introduce the dynamics of the system through the integral form of the above equations and specify the adopted stochastic integral. Throughout all the paper we use the Stratonovich approach for two main reasons: firstly, the usual rules of calculus (for derivatives and integrals) remain unchanged and, secondly, the Stratonovich approach is the most convenient interpretation within the physical sciences since it can be obtained as the limiting process of a coloured noise (with finite noise energy) towards an uncorrelated white one (with diverging energy). Indeed, the white process is a mathematical idealization useful when the typical time-scale of the noise is much shorter than any other time-scale of the system [21].

By the assumption in Eq. (8), we can calculate the generalized forces, obtaining

$$Q_k = -\beta a_{kj}(q)\dot{q}^j + \sum_{i=1}^{N} \sqrt{D m_i}\,\vec{n}_i(t) \cdot \frac{\partial \vec{r}_i}{\partial q^k} - \frac{\partial V}{\partial q^k}. \tag{9}$$

Finally, by introducing the Lagrangian function $\mathcal{L} = T - V$ we can write

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^k} - \frac{\partial \mathcal{L}}{\partial q^k} = -\beta a_{kj}(q)\dot{q}^j + \sum_{i=1}^{N} \sqrt{D m_i}\,\vec{n}_i(t) \cdot \frac{\partial \vec{r}_i}{\partial q^k}, \tag{10}$$

where the right hand side describes the effects of the thermal bath. To conclude, we underline that, being $\partial \vec{r}_i/\partial q^k$ arbitrarily dependent on q, the term $\vec{n}_i(t) \cdot (\partial \vec{r}_i/\partial q^k)$ introduces multiplicative noises in our system. Eq. (10) represents the partial result of this section: a second order Langevin equation with generalized coordinates driven by a field of forces, the friction and the noise.

3 Hamiltonian approach

In order to apply the standard Fokker-Planck methodology we need to work with first order differential equations. To this aim, we transform the Lagrange equations to the Hamilton ones. The explicit form of the Lagrangian function $\mathcal{L}$ is

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2} a_{kh}(q)\dot{q}^k \dot{q}^h - V(q), \tag{11}$$

and the generalized momenta p are defined as $p_h = \partial \mathcal{L}/\partial \dot{q}^h = a_{hj}(q)\dot{q}^j$. Since $\mathbf{A} = \{a_{kh}\}$ is always non-singular [50], we can write $\dot{q}^k = (a^{-1})_{kh} p_h$. Moreover, we can introduce the standard notation $a^{kh} = (a^{-1})_{kh}$, largely used in differential absolute calculus. We therefore obtain the Hamiltonian function as

$$\mathcal{H}(q, p) = p_k \dot{q}^k - \mathcal{L} = \frac{1}{2} a^{kh}(q)\, p_k p_h + V(q). \tag{12}$$

By applying the standard Legendre transformations [50], we eventually obtain the Hamilton equations equivalent to Eq. (10) of the previous section

$$\dot{q}^k = a^{kh} p_h, \tag{13}$$

$$\dot{p}_k = -\frac{1}{2}\frac{\partial a^{ij}}{\partial q^k} p_i p_j - \beta p_k - \frac{\partial V}{\partial q^k} + \sum_{i=1}^{N} \sqrt{D m_i}\,\vec{n}_i(t) \cdot \frac{\partial \vec{r}_i}{\partial q^k}.$$

To conclude, these equations describe an holonomic scleronomous system under the effects of a conservative field and embedded into a thermal bath. From the mathematical point of view, they represent a stochastic differential problem with multiplicative noise (again, with the Stratonovich interpretation).

4 Fokker-Planck methodology

Since we have obtained a first order Langevin differential system, we can directly apply the Fokker-Planck methodology. We summarize here the main result useful to write the Fokker-Planck equation when the Langevin one is given. We remark that the mentioned result is based on the Stratonovitch interpretation of the stochastic differential equations. The following Langevin-like system

$$\frac{dx_i}{dt} = h_i(\vec{x}, t) + \sum_{j=1}^{\mathfrak{m}} g_{ij}(\vec{x}, t) n_j(t) \qquad (\forall i = 1, \ldots, \mathfrak{n}) \tag{14}$$

(with $\mathfrak{n}$ equations and $\mathfrak{m}$ noise terms) corresponds to the following evolution equation for the density probability W (Fokker-Planck equation) [20]

$$\begin{aligned}\frac{\partial W(\vec{x}, t)}{\partial t} = &-\sum_{i=1}^{\mathfrak{n}} \frac{\partial}{\partial x_i}[h_i W(\vec{x}, t)] \\ &-\sum_{i=1}^{\mathfrak{n}} \frac{\partial}{\partial x_i}\left\{\left[\sum_{k=1}^{\mathfrak{n}}\sum_{j=1}^{\mathfrak{m}} g_{kj}\frac{\partial g_{ij}}{\partial x_k}\right] W(\vec{x}, t)\right\} \\ &+\sum_{i=1}^{\mathfrak{n}}\sum_{j=1}^{\mathfrak{m}} \frac{\partial^2}{\partial x_i \partial x_j}\left\{\left[\sum_{k=1}^{\mathfrak{m}} g_{ik}g_{jk}\right] W(\vec{x}, t)\right\},\end{aligned} \tag{15}$$

provided that the Gaussian noises $n_j(t)$ $(\forall j = 1, \ldots, \mathfrak{m})$ satisfy the previously assumed properties: zero average value $(E\{n_j(t)\} = 0)$ and uncorrelation $(E\{n_i(t_1)n_j(t_2)\} = 2\delta_{ij}\delta(t_1 - t_2))$. Note that, with respect to Eqs. (7) and (8), the noise terms in Eq. (14) are now scalars. In Eq. (15), the first term represents the drift term (if $g_{ij} = 0\, \forall i, j$, the equation $\partial W/\partial t = -\sum_{i=1}^{\mathfrak{n}} \partial/\partial x_i(h_i W)$ is in fact the classical Liouville equation), the second one is the noise induced drift term (which is strongly related to the adopted Stratonovitch interpretation, see Refs. [20, 52] for details) and the third one is the diffusion term, describing the effect of the noises. To apply the Fokker-Planck method to our problem, we have to identify Eq. (14) with Eq. (13). To this aim, we can rewrite Eq. (13) as follows

$$\frac{d}{dt}\underbrace{\begin{bmatrix}\vdots\\ q^k\\ \vdots\\ p_k\\ \vdots\end{bmatrix}}_{\vec{x}=(q,p)} = \overbrace{\begin{bmatrix}\vdots\\ a^{kh}p_h\\ \vdots\\ -\frac{1}{2}\frac{\partial a^{ij}}{\partial q^k}p_i p_j - \beta p_k - \frac{\partial V}{\partial q^k}\\ \vdots\end{bmatrix}}^{h_i(\vec{x})} + \underbrace{\begin{bmatrix}0 & \ldots & 0\\ \vdots & \ddots & \vdots\\ 0 & \ldots & 0\\ \sqrt{Dm_1}\frac{\partial x_1}{\partial q^1} & \ldots & \sqrt{Dm_N}\frac{\partial z_N}{\partial q^1}\\ \vdots & \ddots & \vdots\\ \sqrt{Dm_1}\frac{\partial x_1}{\partial q^n} & \ldots & \sqrt{Dm_N}\frac{\partial z_N}{\partial q^n}\end{bmatrix}}_{g_{ij}(\vec{x})}\underbrace{\begin{bmatrix}n_{1x}\\ n_{1y}\\ n_{1z}\\ \vdots\\ n_{Nx}\\ n_{Ny}\\ n_{Nz}\end{bmatrix}}_{n_j(t)}, \tag{16}$$

where it is not difficult to identify the mathematical form of h_i, g_{ij}, n_j, which are the main ingredients of Eq. (14). Of course, we have $\mathfrak{n} = 2n$ and $\mathfrak{m} = 3N$. If we define the matrix $\mathbf{G} = \{g_{ij}\}$, it is easy to recognize that

$$\mathbf{G}\cdot\mathbf{G}^T = \begin{bmatrix}0 & 0\\ 0 & \left\{D\sum_{s=1}^{N} m_i \frac{\partial \vec{r}_s}{\partial q^i}\cdot\frac{\partial \vec{r}_s}{\partial q^j}\right\}\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & \{Da_{ij}\}\end{bmatrix}, \tag{17}$$

where the symbol $\{t_{ij}\}$ represents a matrix with elements t_{ij}. The components of Eq. (17) are easily identified with the diffusion coefficients in Eq. (15). Moreover, it is extremely important to observe, by a simple direct calculation, that the noise induced drift term is exactly zero since $\sum_{k=1}^{\mathfrak{n}=2n}\sum_{j=1}^{\mathfrak{m}=3N} g_{kj}\frac{\partial g_{ij}}{\partial x_k} = 0$. It means that the Itô and the Stratonovich interpretations coincide for Eq. (13) [20]. Summing up, we can write the Fokker-Planck equation in this first form

$$\begin{aligned}\frac{\partial W}{\partial t} = &\frac{1}{2}\frac{\partial W}{\partial p_k}p_i p_j\frac{\partial a^{ij}}{\partial q^k} - a^{kh}\frac{\partial W}{\partial q^k}p_h + n\beta W + \beta p_k\frac{\partial W}{\partial p_k}\\ &+\frac{\partial V}{\partial q^k}\frac{\partial W}{\partial p_k} + Da_{ij}\frac{\partial^2 W}{\partial p_i \partial p_j},\end{aligned} \tag{18}$$

where the implicit summation convention is adopted over the indices running from 1 to n (we will use this notation everywhere). This equation describes the non-equilibrium statistical mechanics of a system represented by generalized coordinates. It is evident that the tensor a_{ij} completely controls the Liouville term (without external fields) and the diffusion one. An intriguing form of this equation can be found by introducing the Poisson brackets as follows

$$\frac{\partial W}{\partial t} = \{\mathcal{H}, W\} + \beta\left\{q^k, p_k W\right\} + Da_{ij}\left\{q^i, \left\{q^j, W\right\}\right\}, \tag{19}$$

where the three terms of drift (Liouville), friction and noise can be easily recognized. This equation can be simply obtained by recalling the definition of Poisson bracket [50]

$$\{f, g\} = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}. \tag{20}$$

As a matter of fact, Eq. (20) allows the determination of the following identities

$$\left\{q^j, W\right\} = \frac{\partial W}{\partial p_j}, \tag{21}$$

$$\left\{q^i, \left\{q^j, W\right\}\right\} = \frac{\partial^2 W}{\partial p_i \partial p_j}, \tag{22}$$

$$\left\{q^k,\, p_k W\right\} = nW + p_k \frac{\partial W}{\partial p_k}, \tag{23}$$

$$\{\mathcal{H},\, W\} = \frac{1}{2}\frac{\partial W}{\partial p_k} p_i p_j \frac{\partial a^{ij}}{\partial q^k} - a^{kh}\frac{\partial W}{\partial q^k} p_h + \frac{\partial V}{\partial q^k}\frac{\partial W}{\partial p_k}, \tag{24}$$

which immediately show the equivalence between Eq. (18) and Eq. (19). Eq. (19) represents a new form of the Klein-Kramers equation for an arbitrary holonomic and scleronomous system embedded in a thermal bath. This is the starting point for the developments presented in next Sections.

It is important to remark that the vector q of the generalized coordinates typically belongs to a well defined subset of the whole n-dimensional space. In other words, $q \in \mathcal{A} \subset \mathbb{R}^n$, where $\mathcal{A}$ describes the configurational space (indeed, with generalized coordinates we may have $\mathcal{A} \neq \mathbb{R}^n$; we can think, e.g., to angles or other non-cartesian variables), while $p \in \mathbb{R}^n$ since there is no limitation on the generalized velocities. Finally, we may write $(q,\, p) \in \mathcal{A} \times \mathbb{R}^n$ with $\mathcal{A} \subset \mathbb{R}^n$.

To conclude, we underline that Eq. (19) can be conveniently modified to write a quantum Klein-Kramers equation, obtained by substituting the density probability W with the reduced density matrix ρ and the Poisson brackets $\{a, b\}$ with the corresponding commutators $\frac{1}{i\hbar}[a, b] = \frac{1}{i\hbar}[ab - ba]$ (correspondence principle) [24]

$$\frac{\partial \rho}{\partial t} = \frac{1}{i\hbar}[\mathcal{H}, \rho] + \frac{\beta}{2i\hbar}\sum_{k=1}^{3}\left[x^k, \Theta_k\rho + \rho\Theta_k\right] - \frac{Dm}{\hbar^2}\sum_{i=1}^{3}\left[x^i, \left[x^i, \rho\right]\right]. \tag{25}$$

Here, for the sake of simplicity, we have considered a single particle of mass m described with orthogonal coordinates x_1, x_2, x_3. Moreover, the product $p_k W$ in Eq. (19) has been substituted by the anti-commutator $\Theta_k\rho + \rho\Theta_k$ between ρ and an operator Θ_k, defined as follows

$$\Theta_k = \frac{1}{2\pi}\iint_{\mathbb{R}^2} e^{i\mathcal{H}\frac{\xi}{\hbar}} p_k e^{-i\mathcal{H}\frac{\xi}{\hbar}} e^{-i\Omega\xi}\, \frac{\tanh\dfrac{\Omega\hbar}{2K_BT}}{\dfrac{\Omega\hbar}{2K_BT}}\, d\xi\, d\Omega. \tag{26}$$

This can be proved by imposing the quantum Gibbs distribution at equilibrium [24]. Even if in Eq. (26) it is quite easy to observe the closeness between Θ_k and p_k, this relation becomes clearer writing the operators in an energy eigenstate basis, as described in Ref. [24]. Finally, the importance of Eq. (19) is not only limited to the classical physics, but it is also relevant for the non-equilibrium quantum statistics.

4.1 Gibbs distribution

The asymptotic behavior of Eq. (18) or (19) is characterized by the following theorem: if the integral defining the partition function

$$Z_{st} = \int_{\mathcal{A}}\int_{\mathbb{R}^n} e^{-\frac{\beta}{D}\mathcal{H}(q,\, p)}\, dq dp \tag{27}$$

is convergent, then the stationary solution of Eq. (18) or (19) is given by the Gibbs distribution in the phase space

$$W_{st}(q,\, p) = \frac{1}{Z_{st}} e^{-\frac{\beta}{D}\mathcal{H}(q,\, p)}. \tag{28}$$

The proof of the latter theorem follows. To begin we observe that

$$\frac{\partial W_{st}}{\partial p_k} = -\frac{\beta}{D} a^{kh} p_h W_{st}, \tag{29}$$

and

$$\frac{\partial W_{st}}{\partial q^k} = -\frac{\beta}{D}\frac{\partial V}{\partial q^k} W_{st} - \frac{1}{2}\frac{\beta}{D}\frac{\partial a^{ij}}{\partial q^k} p_i p_j W_{st}. \tag{30}$$

Moreover,

$$\frac{\partial^2 W_{st}}{\partial p_i \partial p_j} = \frac{\beta^2}{D^2} a^{ih} a^{jk} p_h p_k W_{st} - \frac{\beta}{D} a^{ij} W_{st}. \tag{31}$$

Substituting all terms in Eq. (18) we obtain

$$\begin{aligned}\frac{\partial W_{st}}{\partial t} = &-\frac{1}{2}\frac{\beta}{D} a^{kh} p_h W_{st}\frac{\partial a^{ij}}{\partial q^k} p_i p_j + n\beta W_{st} \\ &- a^{kh} p_h\left[-\frac{\beta}{D}\frac{\partial V}{\partial q^k} W_{st} - \frac{1}{2}\frac{\beta}{D}\frac{\partial a^{ij}}{\partial q^k} p_i p_j W_{st}\right] \\ &- \left(\beta p_k + \frac{\partial V}{\partial q^k}\right)\frac{\beta}{D} a^{kh} p_h W_{st} \\ &+ D a_{ij}\left[\frac{\beta^2}{D^2} a^{ih} a^{jk} p_h p_k W_{st} - \frac{\beta}{D} a^{ij} W_{st}\right],\end{aligned} \tag{32}$$

and, since $\frac{\partial W_{st}}{\partial t} = 0$, $a_{ij}a^{ih} = \delta_j^h$, and $a^{ij}a_{ij} = n$, with straightforward calculations we eventually obtain an identity. This proves that the asymptotic solution is given by the Gibbs distribution, as expected. This result can be further elaborated by considering the explicit expression of the partition function

$$Z_{st} = \int_{\mathcal{A}}\int_{\mathbb{R}^n} e^{-\frac{1}{2}\frac{\beta}{D}a^{kh}p_k p_h} e^{-\frac{\beta}{D}V(q)}\, dq dp, \tag{33}$$

and by using the Gaussian integral

$$\int_{\mathbb{R}^n} e^{-\frac{1}{2}\vec{x}^T\Sigma^{-1}\vec{x}} d\vec{x} = \sqrt{(2\pi)^n \det(\Sigma)}, \tag{34}$$

which is valid for $\vec{x} \in \mathbb{R}^n$ and any symmetric and positive definite matrix Σ [53, 54]. Therefore, we obtain

$$Z_{st} = \left(2\pi \frac{D}{\beta}\right)^{\frac{n}{2}} \int_{\mathcal{A}} \sqrt{\det(\mathbf{A})} e^{-\frac{\beta}{D} V(q)} dq. \tag{35}$$

Now, by means of the Einstein fluctuation-dissipation relation $D = \beta K_B T$ [4, 5] we obtain

$$W_{st}(q, p) = \frac{e^{-\frac{1}{2}\frac{1}{K_B T} p^T \mathbf{A} p} e^{-\frac{1}{K_B T} V(q)}}{\sqrt{(2\pi K_B T)^n} \int_{\mathcal{A}} \sqrt{\det(\mathbf{A})} e^{-\frac{1}{K_B T} V(q)} dq}, \tag{36}$$

where it is important to observe the presence of the term $\sqrt{\det(\mathbf{A})}$, coming from the use of arbitrary non cartesian coordinates. From the phase-space density probability given in Eq. (36) we can obtain the configurational density probability by integration

$$W_c(q) = \int_{\mathbb{R}^n} W_{st}(q, p) dp, \tag{37}$$

or equivalently,

$$W_c(q) = \frac{\int_{\mathbb{R}^n} e^{-\frac{1}{2}\frac{1}{K_B T} p^T \mathbf{A} p} dp \;\; e^{-\frac{1}{K_B T} V(q)}}{\sqrt{(2\pi K_B T)^n} \int_{\mathcal{A}} \sqrt{\det(\mathbf{A})} e^{-\frac{1}{K_B T} V(q)} dq}, \tag{38}$$

and, finally

$$W_c(q) = \frac{\sqrt{\det(\mathbf{A})} e^{-\frac{1}{K_B T} V(q)}}{\int_{\mathcal{A}} \sqrt{\det(\mathbf{A})} e^{-\frac{1}{K_B T} V(q)} dq}. \tag{39}$$

We remark that, while the Gibbs distribution in Eq. (28) represents the classical density probability obtained in statistical mechanics [55–57] (it remains unchanged with the introduction of the generalized coordinates), the reduced configurational counterpart given in Eq. (39) must be written with the essential term $\sqrt{\det(\mathbf{A})}$, which plays a central role in the following developments. We note that it represents the measure element on the differential manifold defined by the metric tensor $\mathbf{A}$.

5 Equivalence of Langevin equations

We have previously shown that the first order differential system describing the non-equilibrium statistical mechanics is given in Eq. (13), with the associated Fokker-Planck (Klein-Kramers) equation given in Eq. (19), being the most important ingredients the Hamiltonian function $\mathcal{H}(q, p)$ and the metric tensor $a_{ij}(q)$. Here, we are interested in investigating the following problem: does some other Langevin system exist with the same associated Klein-Kramers equation? In particular, the following point plays a crucial role: in Eq. (13) we introduced $3N$ noise terms (contained in the N three-dimensional vectors $\vec{n}_1(t),...,\vec{n}_N(t)$) to drive a system with n degrees of freedom. It could be interesting to write a Langevin equation (statistically equivalent to Eq. (13)) with only $n < 3N$ noise terms in order to have a more coherent and elegant theory, and for reducing computational resources if numerical methods are applied. So, we can consider the system

$$\begin{cases} \dot{q}^k = a^{kh} p_h, \\ \dot{p}_k = -\dfrac{1}{2}\dfrac{\partial a^{ij}}{\partial q^k} p_i p_j - \beta p_k - \dfrac{\partial V}{\partial q^k} + \sqrt{D} f_{kj} n^j, \end{cases} \tag{40}$$

where $f_{kj} = f_{kj}(q)$ are unknown functions and $j = 1, ..., n$. It is important to find the form of f_{kj} to have a complete equivalence between Eqs. (13) and (40). It is easy to formulate the problem in terms of the Fokker-Planck equation associated to Eq. (40) and we obtain

$$\frac{\partial W}{\partial t} = \{\mathcal{H}, W\} + \beta \left\{q^k, p_k W\right\} + D f_{ik} f_{jh} \delta^{kh} \left\{q^i, \left\{q^j, W\right\}\right\}. \tag{41}$$

In order to have statistical equivalence we must impose the identity between Eqs. (19) and (41). This leads to

$$f_{ik} f_{jh} \delta^{kh} = a_{ij}. \tag{42}$$

In matrix form we can write $\mathbf{F}\mathbf{F}^T = \mathbf{A}$. To characterize $\mathbf{F}$, we can use the well known polar decomposition theorem (Cauchy) [58]: any non-singular second order tensor $\mathbf{F}$ ($\det(\mathbf{F}) \neq 0$) equals an orthogonal matrix either pre or post multiplied by a positive definite symmetric matrix. Explicitly, we have

$$\forall \mathbf{F} : \det(\mathbf{F}) \neq 0 \Rightarrow \mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}, \tag{43}$$

with $\mathbf{R}^{-1} = \mathbf{R}^T$ and $\mathbf{U}$ and $\mathbf{V}$ symmetric and positive definite. In our case, if we decompose $\mathbf{F}$ as $\mathbf{F} = \mathbf{R}\mathbf{U}$, Eq. (42) assumes the form $\mathbf{R}\mathbf{U}\mathbf{U}^T\mathbf{R}^T = \mathbf{A}$ or, equivalently, $\mathbf{U}^2 = \mathbf{R}^T\mathbf{A}\mathbf{R}$. We can then write $\mathbf{U} = \sqrt{\mathbf{R}^T\mathbf{A}\mathbf{R}} = \mathbf{R}^T\sqrt{\mathbf{A}}\mathbf{R}$, from which we get $\mathbf{F} = \sqrt{\mathbf{A}}\mathbf{R}$. It means that, given $\mathbf{A}$, the tensor $\mathbf{F} = \sqrt{\mathbf{A}}\mathbf{R}$ (for any orthogonal matrix $\mathbf{R}$) fulfils the statistical equivalence. We remark that the square root of any symmetric and positive definite matrix is well defined without ambiguities (in fact, if $\mathbf{A}$ is symmetric and positive definite, by the spectral theorem we have $\mathbf{A} = \mathbf{Q}^T\Lambda\mathbf{Q}$ with $\mathbf{Q}$ orthogonal and Λ diagonal with real positive entries; we then define $\sqrt{\mathbf{A}} = \mathbf{Q}^T\sqrt{\Lambda}\mathbf{Q}$, which is coherent with the usual square root definition).

Finally, the Langevin system

$$\begin{cases} \dot{q}^k = a^{kh} p_h, \\ \dot{p}_k = -\dfrac{1}{2}\dfrac{\partial a^{ij}}{\partial q^k} p_i p_j - \beta p_k - \dfrac{\partial V}{\partial q^k} + \sqrt{D}\left(\sqrt{a}\right)_{kr} R^r_j n^j, \end{cases} \tag{44}$$

is statistically equivalent to Eq. (13) for any orthogonal matrix $\mathbf{R}$. However, the rotation of noises is not really relevant and later on we will use the simplification $\mathbf{R} = \mathbf{I}_n$.

6 The second order Langevin equation

We write here Eq. (44) in form of a second order differential equation. To do this, we simply determine $\ddot{q}^k$ by combining the generalized Hamilton equations

$$\begin{aligned} \ddot{q}^k &= \frac{\partial a^{kh}}{\partial q^i}\dot{q}^i p_h + a^{kh}\dot{p}_h \\ &= \frac{\partial a^{kh}}{\partial q^i} a_{hs}\dot{q}^i\dot{q}^s - \frac{1}{2}a^{kh}\frac{\partial a^{ij}}{\partial q^h} a_{in}a_{jm}\dot{q}^n\dot{q}^m \\ &\quad -\beta a^{kh}a_{hr}\dot{q}^r - a^{kh}\frac{\partial V}{\partial q^h} + \sqrt{D}a^{kh}\left(\sqrt{a}\right)_{hr} R^r_s n^s. \end{aligned} \tag{45}$$

Since $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, by differentiation with respect to q^i we have $\partial\mathbf{A}/\partial q^i\mathbf{A}^{-1} + \mathbf{A}\partial\mathbf{A}^{-1}/\partial q^i = \mathbf{0}$ or, equivalently, $\partial\mathbf{A}^{-1}/\partial q^i = -\mathbf{A}^{-1}\partial\mathbf{A}/\partial q^i\mathbf{A}^{-1}$. In components, we therefore obtain $\frac{\partial a^{kh}}{\partial q^i} = -a^{ks}\frac{\partial a_{st}}{\partial q^i}a^{th}$, which can be used in Eq. (45), eventually yielding

$$\begin{aligned} \ddot{q}^k &= -\frac{1}{2}a^{kh}\left[\frac{\partial a_{hn}}{\partial q^m} + \frac{\partial a_{hm}}{\partial q^n} - \frac{\partial a_{nm}}{\partial q^h}\right]\dot{q}^n\dot{q}^m \\ &\quad -\beta\dot{q}^k - a^{kh}\frac{\partial V}{\partial q^h} + \sqrt{D}a^{kh}\left(\sqrt{a}\right)_{hr} R^r_s n^s. \end{aligned} \tag{46}$$

We can now observe that the definition of the Christoffel symbol of the second kind [59, 60],

$$\left\{\begin{matrix} k \\ nm \end{matrix}\right\} = \frac{1}{2}a^{kh}\left[\frac{\partial a_{hn}}{\partial q^m} + \frac{\partial a_{hm}}{\partial q^n} - \frac{\partial a_{nm}}{\partial q^h}\right], \tag{47}$$

allows us to write Eq. (46) in the simplified form

$$\begin{aligned} &\ddot{q}^k + \left\{\begin{matrix} k \\ nm \end{matrix}\right\}\dot{q}^n\dot{q}^m = \\ &-\beta\dot{q}^k - a^{kh}\frac{\partial V}{\partial q^h} + \sqrt{D}(\sqrt{a^{-1}})^k_r R^r_s n^s, \end{aligned} \tag{48}$$

where we used the identity $a^{kh}(\sqrt{a})_{hr} = (\sqrt{a^{-1}})^k_r$, which means that the element (k, r) of the matrix $\sqrt{\mathbf{A}^{-1}}$ is given by $(\sqrt{a^{-1}})^k_r$.

Eq. (48) is a stochastic differential equation that completely lives on the Riemannian manifold defined by the metric tensor a_{kh}, without the need for a $3N$-dimensional embedding (see Fig. 1 for details). It represents the covariant formulation of the Langevin equation for a system with constraints. The Riemannian manifold has dimension n and, coherently, we have n noise terms. Without friction ($\beta = 0$), drift ($V = 0$) and noise ($D = 0$) the equation assumes the simpler form

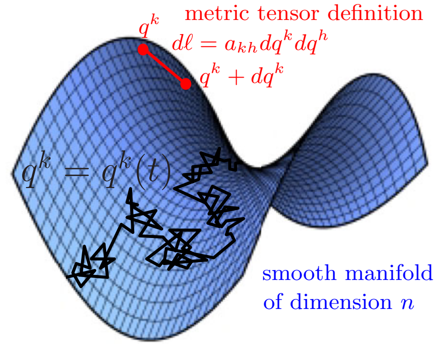

Figure 1 Stochastic dynamics described by Eqs. (44) and (48) on the smooth manifold with metric tensor a_{kh}.

$$\ddot{q}^k + \left\{\begin{matrix} k \\ nm \end{matrix}\right\}\dot{q}^n\dot{q}^m = 0, \tag{49}$$

representing the geodesic curves on the manifold [59, 60]. This case is particularly important since it suggests that the motion of a holonomic scleronomous system (without external actions) can be described by geodesics on a given differential manifold defined by the metric tensor a_{kh}. Therefore, the following analogy holds on: as a free particle follows a straight line in the three-dimensional space, a free holonomic scleronomous system follows a geodetic curve in the pertinent Riemannian manifold. Through Eq. (48) we have generalized this concept by obtaining the covariant representation of the non-equilibrium statistical mechanics (under the Langevin hypotheses): in this general case the system is subjected to a friction and a noise (mimicking the thermal bath) and a drift (induced by the potential energy). Since Eq. (48) is equivalent to Eq. (44), and this latter is equivalent to Eq. (13), the asympototic behavior of Eq. (48) is correctly described (as expected) by the Gibbs distribution. To conclude, we remark that an important point of the previous development is represented by the substitution of the initial $3N$ noise terms with the set of only n random processes in the Langevin approach. Only this technique allows us to obtain a generally covariant equation with a number of noise terms corresponding to the dimension of the manifold (degrees of freedom). It is interesting to remark that Eq. (48), without the potential energy term, is proposed in Refs. [61] as a Langevin equation for covariant state-dependent diffusion processes. Here, differently, it has been exactly derived simply starting from a pure holonomic mechanical system embedded in a classical Langevin thermal bath. Similar approaches can also be found in a work concerning the Brownian motion of systems with stiff bonds [62] and in other investigations conducted in spaces with curvature [63–65] and torsion [66].

7 Analysis of the over-damped system with generalized coordinates

It is well known that the over-damped motion of a particle embedded in a thermal bath and subjected to an external force can be approximately described by a simplified Langevin equation and by the corresponding Fokker-Planck equation, which is the so-called Smoluchowski one. For a non-constrained particle in motion within the three-dimensional space, the exact dynamic equation is

$$m\frac{d^2\vec{r}}{dt^2} = -\frac{\partial V}{\partial \vec{r}} - m\beta\frac{d\vec{r}}{dt} + \sqrt{Dm}\vec{n}, \tag{50}$$

and for large values of β we can write

$$m\frac{d\vec{r}}{dt} = -\frac{1}{\beta}\frac{\partial V}{\partial \vec{r}} + \frac{\sqrt{Dm}}{\beta}\vec{n}, \tag{51}$$

being the inertial term negligible. While the Klein-Kramers equation associated to Eq. (50) exhibits an asymptotic behavior approaching the Gibbs distribution $W_{st}(\vec{r}, \vec{v})$ for large time, the Smoluchowski equation associated to Eq. (51) asymptotically leads to the spatial marginal density $W_c(\vec{r}) = \int_{\Re^3} W_{st}(\vec{r}, \vec{v})d\vec{v}$. Thus, both the Klein-Kramers and the Smoluchowski equations are coherent with standard thermodynamics. Now, the problem is the following: how to find an equation similar to Eq. (51), but describing an arbitrary holonomic scleronomous system? Or, in other terms, how to modify Eq. (48) in order to correctly describe an over-damped system with generalized coordinates? For the moment we simply know that the generally covariant Smoluchowski equation must show the solution given in Eq. (39) for long time. Since the changes to apply to Eq. (48) are not obvious, it is convenient to start from Eq. (39) and to search for a differential system exhibiting an asymptotic behavior given by Eq. (39). We suppose that the first order differential problem corresponding to the over-damped version of Eq. (48) is given by

$$\dot{q}^k = \mu^k(q) + \chi_s^k(q)n^s, \tag{52}$$

where $\mu^k(q)$ and $\chi_s^k(q)$ are unknown functions. Of course, we adopt the Stratonovich interpretation of the stochastic differential equation, as before. Correspondingly, the Smoluchowski equation has the form

$$\frac{\partial W}{\partial t} = -\frac{\partial}{\partial q^k}\left(D^k W\right) + \frac{\partial^2}{\partial q^k \partial q^h}\left(D^{kh}W\right), \tag{53}$$

where drift and diffusion coefficients are given by

$$D^k = \mu^k + \frac{\partial \chi_m^k}{\partial q^h}\chi_n^h\delta^{nm}, \tag{54}$$

$$D^{kh} = \chi_n^k\chi_m^h\delta^{nm}. \tag{55}$$

The asymptotic solution of Eq. (53) must be given by

$$W_c(q) = \frac{1}{Z_c}\sqrt{\det(\mathbf{A})}\exp\left(-\frac{V}{K_B T}\right), \tag{56}$$

where $Z_c = \int_{\mathcal{A}} \sqrt{\det(\mathbf{A})}\exp(-\frac{V}{K_B T})dq$ is the corresponding configurational partition function. To begin, we search for a relation between D^k and D^{kh}, obtained by forcing the solution in Eq. (56) for long time. By differentiating Eq. (56), we have

$$\frac{\partial W_c(q)}{\partial q^j} = \left(\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial\sqrt{\det(\mathbf{A})}}{\partial q^j} - \frac{1}{K_B T}\frac{\partial V}{\partial q^j}\right)W_c. \tag{57}$$

By substituting W_c in Eq. (53) we have $0 = -D^k W_c + \partial/\partial q^h(D^{kh}W_c)$ and by using Eq. (57) we simply get

$$D^k = \frac{\partial D^{kh}}{\partial q^h} + \frac{D^{kh}}{\sqrt{\det(\mathbf{A})}}\frac{\partial\sqrt{\det(\mathbf{A})}}{\partial q^h} - \frac{D^{kh}}{K_B T}\frac{\partial V}{\partial q^h}. \tag{58}$$

This is an explicit relation between D^k and D^{kh}, which must be verified for the validity of the Smoluchowski equation. It can be further elaborated by using the rule for the derivative of a determinant

$$\frac{\partial \det(\mathbf{A})}{\partial q^j} = \det(\mathbf{A})\mathrm{tr}\left(\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial q^j}\right), \tag{59}$$

which holds for any non-singular matrix $\mathbf{A}$. We can also calculate

$$\frac{\partial\sqrt{\det(\mathbf{A})}}{\partial q^j} = \frac{\sqrt{\det(\mathbf{A})}}{2}a^{kh}\frac{\partial a_{hk}}{\partial q^j}, \tag{60}$$

and, therefore, we obtain from Eqs. (58), (54) and (55)

$$\mu^k = \chi_n^k\frac{\partial\chi_m^h}{\partial q^h}\delta^{nm} + \frac{1}{2}\chi_n^k\chi_m^h\delta^{nm}a^{rt}\frac{\partial a_{rt}}{\partial q^h} - \frac{1}{K_B T}\chi_n^k\chi_m^h\delta^{nm}\frac{\partial V}{\partial q^h}. \tag{61}$$

This is an exact relation between μ^i and χ_j^i imposing the correct asymptotic behavior in Eqs. (52) and (53). In particular, Eq. (52) can now be written as

$$\dot{q}^k = \chi_n^k \frac{\partial \chi_m^h}{\partial q^h}\delta^{nm} + \frac{1}{2}\chi_n^k\chi_m^h\delta^{nm}a^{rt}\frac{\partial a_{rt}}{\partial q^h} - \frac{1}{K_B T}\chi_n^k\chi_m^h\delta^{nm}\frac{\partial V}{\partial q^h} + \chi_s^k n^s. \tag{62}$$

We have to compare this over-damped dynamics with Eq. (48) divided by β where, for the sake of simplicity, we impose $R_s^r = \delta_s^r$ (indeed, by cancelling out the effect of rotations we do not lose the generality)

$$\frac{1}{\beta}\ddot{q}^k + \frac{1}{\beta}\left\{ \begin{matrix} k \\ nm \end{matrix} \right\}\dot{q}^n\dot{q}^m = - \dot{q}^k - \frac{1}{\beta}a^{kh}\frac{\partial V}{\partial q^h} + \sqrt{\frac{K_B T}{\beta}}(\sqrt{a^{-1}})_s^k n^s. \tag{63}$$

We observe the same friction term in Eqs. (62) and (63) and we impose the same mathematical form for the potential energy and noise terms, eventually getting

$$\chi_n^k\chi_m^h\delta^{nm} = \frac{K_B T}{\beta}a^{kh} \text{ and } \chi_s^k = \sqrt{\frac{K_B T}{\beta}}(\sqrt{a^{-1}})_s^k. \tag{64}$$

Interestingly, the identities coming from the potential energy and the noise terms are exactly the same, as shown in Eq. (64). This result completes the determination of the balance equations describing the over-damped motion of the system. Indeed, by using Eqs. (61) and (64), all necessary terms to define the over-damped Langevin equation and the corresponding Smoluchowski equation are fully known. However, their mathematical form can be strongly simplified in order to better show the covariant character of these equations.

7.1 Covariant over-damped Langevin equation

To begin, we try to write Eq. (58) in a more convenient form. By performing the following transformations

$$\begin{aligned} D^k &= \frac{\partial}{\partial q^h}\frac{D^{kh}\sqrt{\det(\mathbf{A})}}{\sqrt{\det(\mathbf{A})}} + \frac{D^{kh}}{\sqrt{\det(\mathbf{A})}}\frac{\partial\sqrt{\det(\mathbf{A})}}{\partial q^h} - \frac{D^{kh}}{K_B T}\frac{\partial V}{\partial q^h} \\ &= \frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^h}\left(D^{kh}\sqrt{\det(\mathbf{A})}\right) \\ &\quad + D^{kh}\sqrt{\det(\mathbf{A})}\frac{\partial}{\partial q^h}\frac{1}{\sqrt{\det(\mathbf{A})}} \\ &\quad + \frac{D^{kh}}{\sqrt{\det(\mathbf{A})}}\frac{\partial\sqrt{\det(\mathbf{A})}}{\partial q^h} - \frac{D^{kh}}{K_B T}\frac{\partial V}{\partial q^h}, \end{aligned} \tag{65}$$

it is not difficult to prove that the second and third terms cancel out, eventually obtaining

$$D^k = \frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^h}\left(D^{kh}\sqrt{\det(\mathbf{A})}\right) - \frac{D^{kh}}{K_B T}\frac{\partial V}{\partial q^h}. \tag{66}$$

It is now useful to recall some results concerning the covariant derivatives on a differential manifold [59, 60]. The covariant derivative of a vector (expressed in controvariant components) is given by

$$T^i_{\|s} = \frac{\partial T^i}{\partial q^s} + \left\{ \begin{matrix} i \\ sk \end{matrix} \right\} T^k \tag{67}$$

where $T^i_{\|s}$ means covariant derivative of T^i with respect to q^s and, as before, $\left\{ \begin{matrix} i \\ sk \end{matrix} \right\}$ represents the Chrystoffel symbol associated to the metric tensor a_{ij}. The covariant divergence is therefore defined as

$$T^h_{\|h} = \frac{\partial T^h}{\partial q^h} + \left\{ \begin{matrix} h \\ hk \end{matrix} \right\} T^k. \tag{68}$$

Classical transformations, well-known in absolute differential calculus [60], prove that

$$T^h_{\|h} = \frac{\partial T^h}{\partial q^h} + \frac{1}{2\det(\mathbf{A})}\frac{\partial \det(\mathbf{A})}{\partial q^k}T^k, \tag{69}$$

or, equivalently,

$$T^h_{\|h} = \frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^h}\left(T^h\sqrt{\det(\mathbf{A})}\right). \tag{70}$$

Therefore, the first term in the right hand side of Eq. (66) represents the divergence of D^{kh} with respect to the second index. Now, since $D^{kh} = \chi_n^k\chi_m^h\delta^{nm} = \frac{K_B T}{\beta}a^{kh}$ (see Eqs. (55) and (64)) we have

$$D^k = \frac{K_B T}{\beta\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^h}\left(a^{kh}\sqrt{\det(\mathbf{A})}\right) - \frac{a^{kh}}{\beta}\frac{\partial V}{\partial q^h}, \tag{71}$$

the first term being now proportional to the covariant divergence of a^{kh}. Hence, we can write

$$D^k = \frac{K_B T}{\beta}a^{kh}_{\|h} - \frac{D^{kh}}{K_B T}\frac{\partial V}{\partial q^h}, \tag{72}$$

and, by using the divergence definition given in Eq. (68), we can find

$$\begin{aligned} a^{kh}_{\parallel h} &= \frac{\partial a^{kh}}{\partial q^h} + \left\{ \begin{matrix} h \\ hs \end{matrix} \right\} a^{ks} \\ &= -a^{ks} \frac{\partial a_{st}}{\partial q^h} a^{th} + \frac{1}{2} a^{th} \left[\frac{\partial a_{th}}{\partial q^s} + \frac{\partial a_{ts}}{\partial q^h} - \frac{\partial a_{hs}}{\partial q^t} \right] a^{ks} \\ &= -a^{ks} a^{th} \left[\frac{\partial a_{st}}{\partial q^h} - \frac{1}{2} \frac{\partial a_{th}}{\partial q^s} \right] \\ &= -a^{ks} a^{th} \frac{1}{2} \left[\frac{\partial a_{st}}{\partial q^h} + \frac{\partial a_{sh}}{\partial q^t} - \frac{\partial a_{th}}{\partial q^s} \right] \\ &= -a^{th} \left\{ \begin{matrix} k \\ th \end{matrix} \right\}. \end{aligned} \tag{73}$$

Summing up, Eq. (72) assumes the important form

$$D^k = -\frac{K_B T}{\beta} a^{st} \left\{ \begin{matrix} k \\ st \end{matrix} \right\} - \frac{a^{kh}}{\beta} \frac{\partial V}{\partial q^h}, \tag{74}$$

and the Langevin over-damped equation can be written as

$$\begin{aligned} \dot{q}^k = &-\frac{K_B T}{\beta} a^{st} \left\{ \begin{matrix} k \\ st \end{matrix} \right\} - \frac{a^{kh}}{\beta} \frac{\partial V}{\partial q^h} \\ &- \frac{\partial \chi^k_m}{\partial q^h} \chi^h_n \delta^{nm} + \chi^k_s n^s. \end{aligned} \tag{75}$$

where we used Eq. (54) to obtain the drift field μ^k. In Eq. (75) the coefficients $\chi^k_s = \sqrt{\frac{K_B T}{\beta}} (\sqrt{a^{-1}})^k_s$ must be taken into consideration, as stated in Eq. (64). This first order differential system represents the over-damped version of Eq. (48), describing the stochastic motion of the system. As before, Eq. (75) is written by considering the Stratonovich interpretation. However, it is interesting to observe that its form is even simpler if we adopt the Itô formalism. Indeed, we simply have

$$\dot{q}^k \overset{(\text{Itô})}{=} -\frac{K_B T}{\beta} a^{st} \left\{ \begin{matrix} k \\ st \end{matrix} \right\} - \frac{a^{kh}}{\beta} \frac{\partial V}{\partial q^h} + \sqrt{\frac{K_B T}{\beta}} (\sqrt{a^{-1}})^k_s n^s. \tag{76}$$

Here, to transform Eq. (75) into Eq. (76) we used the standard transformation rules between the Itô and the Stratonovich formalisms [20, 21, 51, 52]. In particular, if we have a stochastic differential equation of the form given in Eq. (14), interpreted with the Stratonovich formalism, we can write the equivalent equation in the Itô formalism as follows

$$\frac{dx_i}{dt} \overset{(\text{Itô})}{=} \hat{h}_i(\vec{x}, t) + \sum_{j=1}^{m} g_{ij}(\vec{x}, t) n_j(t),$$

where

$$\hat{h}_i(\vec{x}, t) = h_i(\vec{x}, t) + \sum_{k=1}^{n} \sum_{j=1}^{m} \frac{\partial g_{ij}(\vec{x}, t)}{\partial x_k} g_{kj}(\vec{x}, t) \tag{77}$$

represents the new drift coefficient of the Itô stochastic differential equation (the added term is sometimes called Wong-Zakai correction, as described in Chapter 3 of Ref. [52]). This property has been used to derive Eq. (76) from Eq. (75).

Moreover, the free motion over the manifold ($V = 0$) within the Itô formalism assumes the form

$$\dot{q}^k \overset{(\text{Itô})}{=} -\frac{K_B T}{\beta} a^{st} \left\{ \begin{matrix} k \\ st \end{matrix} \right\} + \sqrt{\frac{K_B T}{\beta}} (\sqrt{a^{-1}})^k_s n^s. \tag{78}$$

It is interesting to note that this equation is typically used to define the Brownian motion on a Riemannian manifold within the mathematical community [67–70]. We remark that, in this work, we reobtain this equation as the result of the application of the analytical and statistical mechanics to a constrained and over-damped system described through generalized coordinates. Thus, the classical mathematical definition of the Brownian motion on a manifold exactly corresponds to the over-damped version of the non-equilibrium statistical mechanics (introduced through the Langevin-Fokker-Planck methodology).

The obtained Langevin equations for a constrained system (see Eq. (44) or (48) for the motion in the whole phase space or Eq. (75) for its over-damped version) can be adopted to develop an *ad hoc* Langevin molecular dynamics able to analyse the time behavior of holonomic structures. This is an important topic in computational physics, as largely described in recent literature [71, 72]. We underline that for the implementation of Eq. (44), (48) or (75) through a standard discretization over the time, we need to calculate the square root of the metric tensor at any time step (see, for instance, Eq. (64)). This can be done with classical numerical approximations [73, 74]. The important point is the following: by means of our approach, the problem of the constraints, crucial for many numerical applications, is automatically solved and it does not require any other particular expedient. Moreover, the convergence properties of finite difference schemes for Stratonovich stochastic differential equations are largely studied in literature [51, 52]. We also remark that Eq. (76) can be directly implemented through specific numerical schemes expressly devoted to Itô stochastic differential equations [75, 76]. Being the aim of this paper mainly theoretical, the comparison between the proposed Langevin molecular dynamics and earlier techniques is left for future investigations. Anyway, these methodologies play a central role in studying the out of equilibrium response of constrained molecular structures such as chains with rigid bonds or polymers confined over arbitrary surfaces.

7.2 Covariant Smoluchowski equation

To conclude, we also study here the covariant form of the Smoluchowski equation introduced in Eq. (53). Instead of analysing the dynamics of $W(q, t)$, it is preferable to consider the associated density [77–79]

$$P(q, t) = \frac{1}{\sqrt{\det(\mathbf{A})}} W(q, t), \tag{79}$$

which is more indicated to obtain a generally covariant version of Eq. (53). Since $W(q, t)$ satisfies Eq. (53), the quantity $P(q, t)$ is governed by the following equation

$$\sqrt{\det(\mathbf{A})}\,\frac{\partial P}{\partial t} = -\frac{\partial}{\partial q^k}\left(D^k\sqrt{\det(\mathbf{A})}\,P\right) + \frac{\partial^2}{\partial q^k \partial q^h}\left(D^{kh}\sqrt{\det(\mathbf{A})}\,P\right). \tag{80}$$

The second term in Eq. (80) can be elaborated as follows

$$\begin{aligned} &\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial^2}{\partial q^k \partial q^h}\left(\frac{K_B T}{\beta} a^{kh}\sqrt{\det(\mathbf{A})}\,P\right) \\ &= \frac{K_B T}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left[\frac{\partial}{\partial q^h}\left(a^{kh}\sqrt{\det(\mathbf{A})}\right)P\right] \\ &\quad + \frac{K_B T}{\beta}\nabla^2 P \end{aligned} \tag{81}$$

where ∇^2 represents the Laplace-Beltrami operator (the equivalent of the standard Laplace operator on a differential manifold), defined as follows [60]

$$\nabla^2 P = \frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left(\sqrt{\det(\mathbf{A})}\,a^{kh}\frac{\partial P}{\partial q^h}\right). \tag{82}$$

Similarly, the first term in Eq. (80) can be modified by using Eq. (66), by obtaining

$$\begin{aligned} &-\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left(D^k\sqrt{\det(\mathbf{A})}\,P\right) \\ &= -\frac{K_B T}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left[\frac{\partial}{\partial q^h}\left(a^{kh}\sqrt{\det(\mathbf{A})}\right)P\right] \\ &\quad + \frac{1}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left[a^{kh}\frac{\partial V}{\partial q^h}\sqrt{\det(\mathbf{A})}\,P\right]. \end{aligned} \tag{83}$$

Summing up, we have the final equation

$$\frac{\partial P}{\partial t} = \frac{1}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left(a^{kh}\frac{\partial V}{\partial q^h}\sqrt{\det(\mathbf{A})}\,P\right) + \frac{K_B T}{\beta}\nabla^2 P, \tag{84}$$

which is the generally covariant form of the Smoluchowski equation. The first drift term is represented by the covariant divergence of $a^{kh}\partial V/\partial q^h P$ and describes the effects of the external forces. The second diffusion term corresponds to the Laplace-Beltrami operator applied to P. By observing the explicit form of Eq. (84), i.e.

$$\begin{aligned} \frac{\partial P}{\partial t} &= \frac{1}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left(a^{kh}\frac{\partial V}{\partial q^h}\sqrt{\det(\mathbf{A})}\,P\right) \\ &\quad + \frac{K_B T}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left(a^{kh}\frac{\partial P}{\partial q^h}\sqrt{\det(\mathbf{A})}\right), \end{aligned} \tag{85}$$

we can define an effective total potential ϕ

$$\phi = V + K_B T \log P, \tag{86}$$

composed of the mechanical potential energy V and the chemical potential $K_B T \log P$. Hence, Eq. (85) assumes the simpler form

$$\frac{\partial P}{\partial t} = \frac{1}{\beta}\frac{1}{\sqrt{\det(\mathbf{A})}}\frac{\partial}{\partial q^k}\left(a^{kh}\frac{\partial \phi}{\partial q^h}\sqrt{\det(\mathbf{A})}\,P\right), \tag{87}$$

which can be interpreted by stating that both the drift component and the diffusion one can be introduced by means of a single effective potential function given in Eq. (86). In other terms, we can affirm that the purely mechanical force $f_h^{drift} = -\partial V/\partial q^h$ (which is valid for $T = 0$) must be substituted with the total force $f_h^{tot} = f_h^{drift} + f_h^{diff}$, where $f_h^{diff} = -K_B T \partial \log P/\partial q^h$, when the system is embedded in a thermal bath with a finite temperature $T > 0$. Moreover, the diffusion term in previous expression is at the origin of the Fick's law of diffusion [3, 80]. In fact, through Eq. (87) written in the continuity form $\partial P/\partial t + J^k_{\parallel k} = 0$, we can identify the contravariant flux $J^k = -(1/\beta) P a^{kh}\partial\phi/\partial q^h$ or, equivalently, its covariant version $J_k = -(1/\beta) P \partial\phi/\partial q^k$. This flux can be written in explicit form as

$$J_k = -\frac{P}{\beta}\frac{\partial \phi}{\partial q^k} = -\frac{P}{\beta}\frac{\partial V}{\partial q^k} - \frac{K_B T}{\beta}\frac{\partial P}{\partial q^k}, \tag{88}$$

where the first term corresponds to the drift flux (being $-(1/\beta)\partial V/\partial q^k$ the k–th component of the asymptotic velocity and P equivalent to the concentration) and the second one represents the classical Fick's flux, proportional to the negative gradient of concentration.

To conclude, if we consider the evolution of a free system ($V = 0$), the Smoluchowski equation collapses into the standard heat equation on a manifold

$$\frac{\partial P}{\partial t} = \frac{K_B T}{\beta} \nabla^2 P. \tag{89}$$

In this particular case the asymptotic solution is given by

$$\lim_{t \to \infty} P(q, t) = \frac{1}{\int_{\mathcal{A}} \sqrt{\det(\mathbf{A})} dq}, \tag{90}$$

representing the uniform distribution on the manifold ($q \in \mathcal{A}$).

8 Conclusions

In this work we developed a thorough formalism concerning the effects of a Langevin bath on a mechanical holonomic system. By generalizing the classical Lagrange and Hamilton equations with suitable friction and noise terms, we work out a non-equilibrium statistical theory working with an arbitrary set of generalized coordinates. If we frame the analysis within the complete phase space, we achieve the Langevin and Klein-Kramers equations (see Eqs. (19) and (44)) based on a set of noises containing as many terms as the degrees of freedom n of the system. So doing, these equations completely live within the smooth manifold of dimension n, defined by the metric tensor of the system. On the other hand, if we consider the reduced configurational space, our construction yields the generally covariant form of the Langevin and Smoluchowski equations (see Eqs. (75) and (84)). These results can be naturally interpreted by concluding that the non-equilibrium evolution of the system corresponds to a Brownian motion on a Riemannian manifold driven by the externally applied forces. Indeed, if the system is free (not subjected to any external potential energy), the final equations exactly correspond to the definition of the Brownian motion on a manifold, as postulated within the mathematical community. It is also interesting to remark that the obtained form of the Smoluchowski equation allows us to immediately deduce the diffusion Fick's law in a generally covariant setting. The knowledge of the exact equations governing the non-equilibrium statistical mechanics of constrained systems is of central importance for several applications ranging from polymer science to soft matter or biological physics. In particular, our results can be used to develop molecular dynamics methods for constrained structures, advantageous for studying assemblages with stiff bonds and rigid constraints.

Key words. Non-equilibrium statistical mechanics, Langevin and Fokker-Planck equations, Smoluchowski equation, differential absolute calculus, smooth manifolds.

References

[1] R. Brown, Philos. Mag. N. S. **4**, 161 (1828).
[2] R. Brown, Philos. Mag. N. S. **6**, 161 (1829).
[3] A. Fick, Ann. Phys. (Leipzig) **19**, 371 (1855).
[4] A. Einstein, Ann. Phys. (Leipzig) **17**, 549 (1905).
[5] A. Einstein, Ann. Phys. (Leipzig) **19**, 371 (1906)
[6] M. von Smoluchowski, Ann. Phys. (Leipzig) **21**, 756 (1906).
[7] J. B. Perrin, C. R. Acad. Sci. (France) **158**, 1168 (1914).
[8] J. B. Perrin, Discontinuous Structure of Matter, Nobel Lecture (1926).
[9] P. Langevin, C. R. Acad. Sci. (France) **146**, 530 (1908).
[10] A. D. Fokker, Ann. Phys. **348**, 810 (1914).
[11] M. Planck, Sitzber. Preuss. Akad. Wiss. **324** (1917).
[12] A. Kolmogoroff, Mathematische Annalen **104**, 415 (1931).
[13] O. Klein, Arkiv für Matematik, Astronomi och Fysik **16**, 1 (1921).
[14] H. A. Kramers, Physica **7**, 284 (1940).
[15] G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).
[16] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945).
[17] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[18] W. Coffey, Advances in Chemical Physics, Vol.LXIII (1985).
[19] N. G. Van Kampen, Stochastic Processes in Physics and Chemistry (Elsevier, Amsterdam, 1981).
[20] H. Risken, The Fokker-Planck equation (Springer Verlag, Berlin, 1989).
[21] W. T. Coffey, Yu. P. Kalmykov, and J. P. Waldron, The Langevin equation (World Scientific, Singapore, 2004).
[22] P. Caldirola and L. A. Lugiato, Physica A **116**, 248 (1982).
[23] A. O. Caldeira and A. J. Leggett, Physica A **121**, 587 (1983).
[24] B. Bianco, E. Moggia, S. Giordano, W. Rocchia, and A. Chiabrera, Il Nuovo Cimento **116**, 155 (2001).
[25] P. Pechukas, J. Ankerhold, and H. Grabert, Ann. Phys. (Leipzig) **9**, 794 (2000).
[26] E. Frey and K. Kroy, Ann. Phys. (Leipzig) **14**, 20 (2005).
[27] M. Doi, Introduction to Polymer Physics (Clarendon Press, Oxford, 1996).
[28] M. Rubinstein and R. H. Colby, Polymer Physics (Oxford University Press, New York, 2003).
[29] S. M. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).

[30] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995).
[31] F. Ritort, J. Phys.: Condens. Matter **18**, R531 (2006).
[32] F. Manca, S. Giordano, P. L. Palla, R. Zucca, F. Cleri, and L. Colombo, J. Chem. Phys. **136**, 154906 (2012).
[33] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo, J. Chem. Phys. **137**, 244907 (2012).
[34] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo, J. Phys.: Conf. Ser. **383**, 012016 (2012).
[35] M. Rief, J. M. Fernandez, and H. E. Gaub, Phys. Rev. Lett. **81**, 4764 (1998).
[36] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo, Phys. Rev. E **87**, 032705 (2013).
[37] D. De Tommasi, N. Millardi, G. Puglisi, and G. Saccomandi, J. R. Soc. Interface **10**, 20130651 (2013).
[38] F. Manca, S. Giordano, P. L. Palla, and F. Cleri, Physica A: Statistical Mechanics and its Applications **395**, 154 (2014).
[39] K. Svoboda, C. Schmidt, B. Schnapp, and S. Block, Nature **365**, 721 (1993).
[40] M. Magnasco, Phys. Rev. Lett. **71**, 1477 (1993).
[41] L. Landau and E. Lifshitz, Phys. Z. Sowjetunion **8**, 153 (1935).
[42] T. L. Gilbert, Phys. Rev. **100**, 1243 (1955) (abstract only); IEEE Trans. Mag. 40, 3443 (2004).
[43] W. F. Brown, J. Appl. Phys. **30**, S130 (1959).
[44] W. F. Brown, IEEE Trans. Magn. **15**, 1196 (1979).
[45] G. Bertotti, I. Mayergoyz, and C. Serpico, Nonlinear Magnetization Dynamic in Nanosystems (Elsevier, Oxford, 2000).
[46] D. R. Fredkin, Physica B **306**, 26 (2001).
[47] P. M. Déjardin, D. S. F. Crothers, W. T. Coffey, and D. J. McCarthy, Phys. Rev. E **63**, 021102 (2001).
[48] Y. P. Kalmykov, W. T. Coffey, U. Atxitia, O. Chubykalo-Fesenko, P. M. Déjardin, and R. W. Chantrell, Phys. Rev. B **82**, 024412 (2010).
[49] S. Giordano, Y. Dusch, N. Tiercelin, P. Pernod, and V. Preobrazhensky, Eur. Phys. J. B **86**, 249 (2013).
[50] F. Gantmacher, Lectures in Analytical Mechanics (MIR Publisher, Moscou, 1975).
[51] Z. Schuss, Theory and Applications of Stochastic Differential Equations (John Wiley & Sons, New York, 1980).
[52] Z. Schuss, Theory and Applications of Stochastic Processes (Springer, New York, 2010).
[53] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products (Academic Press, San Diego, 1965).
[54] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions (Dover Publication, New York, 1970).
[55] J. H. Weiner, Statistical mechanics of elasticity (Dover Publication Inc., New York, 2002).
[56] D. Chandler, Introduction to Modern Statistical Mechanics (Oxford University Press, New York, 1987).
[57] J. W. Gibbs, Elementary principles in statistical mechanics (Charles Scribner's Sons, New York, 1902).
[58] R. J. Atkin and N. Fox, An introduction to the theory of elasticity (Dover Publication Inc., New York, 1980).
[59] M. P. do Carmo, Differential Geometry of Curves and Surfaces (Prentice-Hall, New York, 1976).
[60] T. Levi-Civita, The absolute differential calculus (Blackie & Son Limited, London, 1946).
[61] M. Polettini, J. Stat. Mech.: Theory and Exper., P07005 (2013).
[62] E. J. Hinch, J. Fluid Mech. **271**, 219 (1994).
[63] P. Castro-Villarreal, J. Stat. Mech.: Theory and Exper., P08006 (2010) .
[64] P. Castro-Villarreal, J. Stat. Mech.: Theory and Exper., P05017 (2014).
[65] P. Castro-Villarreal, A. Villada-Balbuena, J. M. Méndez-Alcaraz, R. Casta neda-Priego, and S. Estrada-Jiménez, J. Chem. Phys. **140**, 214115 (2014).
[66] H. Kleinert and S. V. Shabanov, J. Phys. A: Math. Gen. **31**, 7005 (1998).
[67] P. Hsu, Contemporary mathematics AMS **73**, 95 (1988).
[68] K. Itô, Proc. Internat. Congr. Math., Stockholm, Inst. Mittag-Leffler, Djursholm, p. 536-539 (1962).
[69] W. S. Kendall, Acta Applicandae Mathematicae **9**, 29 (1987).
[70] D. W. Stroock, Bull. AMS **33**, 135 (1996).
[71] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden, ChemPhysChem **6**, 1809 (2005).
[72] T. Lelièvre, M. Rousset, and G. Stolz, Mathematics of computation **81**, 2071 (2012).
[73] N. J. Higham, Linear Algebra and its Applications **88**, 405 (1987).
[74] E. J. Allen, J. Baglama, and S. K. Boyd, Linear Algebra and its Applications **310**, 167 (2000).
[75] G. N. Mil'shtejn, Theory of Probability & Its Applications **19**, 557 (1975).
[76] T. Sauer, Interdisciplinary Reviews: Computational Statistics **5**, 362 (2013).
[77] R. Graham, Z. Physik B **26**, 397 (1977).
[78] H. Grabert and M. S. Green, Phys. Rev. A **19**, 1747 (1979).
[79] H. Grabert, R. Graham, and M. S. Green, Phys. Rev. A **21**, 2136 (1980).
[80] J. Crank, The Mathematics of Diffusion (Oxford University Press, Oxford, 1980).