Generally covariant state-dependent diffusion

Matteo Polettini

Complex Systems and Statistical Mechanics, University of Luxembourg, Campus Limpertsberg, 162a avenue de la Faëncerie, L-1511 Luxembourg, Luxembourg
E-mail: matteo.polettini@uni.lu

Received 3 December 2012
Accepted 9 June 2013
Published 16 July 2013

Online at stacks.iop.org/JSTAT/2013/P07005
doi:10.1088/1742-5468/2013/07/P07005

Abstract. Statistical invariance of Wiener increments under $SO(n)$ rotations provides a notion of gauge transformation of state-dependent Brownian motion. We show that the stochastic dynamics of non-gauge-invariant systems is not unambiguously defined. They typically do not relax to equilibrium steady states even in the absence of external forces. Assuming both coordinate covariance and gauge invariance, we derive a second-order Langevin equation with state-dependent diffusion matrix and vanishing environmental forces. It differs from previous proposals but nevertheless incorporates the Einstein relation, a Maxwellian conditional steady state for the velocities, and the equipartition theorem. The overdamping limit leads to a stochastic differential equation in state space that cannot be interpreted as a pure differential (Itô, Stratonovich or other). At odds with the latter interpretations, the corresponding Fokker–Planck equation admits an equilibrium steady state; a detailed comparison with other theories of state-dependent diffusion is carried out. We propose this as a theory of diffusion in a heat bath with varying temperature. Besides equilibrium, a crucial experimental signature is the nonuniform steady spatial distribution.

Keywords: stationary states, diffusion

ArXiv ePrint: 1206.2798
## Contents

1. Introduction 3  
   1.1. Scope and motivation 3  
   1.2. Models and results 3  
   1.3. Plan and notation 5  

2. Brownian motion 5  
   2.1. Gauge invariance under rotations 5  
   2.2. Generally covariant Langevin equation 7  
   2.3. Kramers' equation and its steady state 9  

3. Overdamped dynamics 10  
   3.1. Fokker–Planck equation 10  
   3.2. Smoluchovski equation 10  
   3.3. Covariant entropy and entropy production 11  

4. Geometric insights 12  
   4.1. Gauge connection 12  
   4.2. Nonparallelizable manifolds: the sphere 14  
   4.3. Redundant and defective sources of noise 14  

5. Comparison of models 15  
   5.1. (B) Second-order Graham theory 16  
   5.2. (C) State-dependent temperature 16  
   5.3. (G) Metric/diffusion theory 17  
   5.4. (D) State-dependent viscosity 17  
   5.5. (E) Pure posticipatory SDE 18  
   5.6. (F) State-dependent viscosity and temperature 18  

6. Conclusions 18  

Acknowledgments 19  

Appendix A. Exponential decay to equipartition 20  

Appendix B. Overdamping limit of metric/diffusion theory 20  

Appendix C. Overdamping limit of diffusion with state-dependent temperature and viscosity 22  

Appendix D. Comparison of models 23  

References 26
1. Introduction

Nonequilibrium thermodynamics is often modeled by means of stochastic differential equations (SDEs), describing possible trajectories in an open system’s state space, and by diffusion equations for the evolution of ensembles [1]–[3]. While the paradigm started by enhancing dynamical equations with an external source of noise, independent of the system’s state, soon there arose the necessity of considering situations where the intensity and direction of the environmental disturbance depends on the state of the system [4]–[6]. From a different perspective, mathematicians had extensively studied SDEs on manifolds, where a state-dependent diffusion coefficient is often associated to a Riemannian metric (see [7, 8] for recent monographs). At the merge between the two approaches Van Kampen argued that ‘The question of the existence and correct form of equations describing Brownian motion on a manifold cannot be answered by mathematics alone, but requires a study of the underlying physics’ [9].

1.1. Scope and motivation

While statistical properties of Brownian motion are expected to only depend on the diffusion matrix and the physical drift—loosely speaking the ‘square root’ of the diffusion matrix appears in a SDE. In one dimension this is unambiguous, while in $n$ dimensions there are many different ways to take such a square root, differing one from another by local internal rotations. We will require the physical description to be independent of this choice. When this does not occur, for a given diffusion matrix one can write down SDEs that have very different dynamics and thermodynamics. In particular, most choices lead to nonequilibrium steady states even in the absence of external forces: an unobservable statistical property oddly becomes a nonequilibrium drive.

Taking up Van Kampen’s invitation, in this paper we make a proposal for state-dependent diffusion equations based on the physical symmetry principle of general covariance, by which we mean (1) covariance under external coordinate transformations and (2) invariance under local internal rotations (gauge invariance). The covariance of nonequilibrium statistical mechanics under coordinate transformations has been extensively studied by Graham [10] (see also [11] for applications to semiclassical methods in field theory). Departing from Graham’s analysis, we further assume that statistical properties of the theory are invariant under a class of internal gauge transformations under rotations of the noise components, taking inspiration from the frame field approach to general relativity. However, we do not derive relativistic equations for diffusion in strong gravitational fields.

Recently there has been an increasing interest in geometric methods in nonequilibrium statistical mechanics [12]–[17], with emphasis often given to the equilibrium versus nonequilibrium character of steady states. This paradigm will allow us to better appreciate the physical nature of our construction. It is our conviction that the theory is relevant to diffusion in environments with nonuniform temperature, as we anticipate in the next paragraph and argue more thoroughly in a parallel publication [18].

1.2. Models and results

In this paper we consider stochastic models of systems subject to state-dependent diffusion, viscous damping and no external forces, unless otherwise stated. We do not
Generally covariant state-dependent diffusion

consider velocity-dependent diffusion, as is done in [6]. For this, second-order Langevin-type equations serve the purpose better than do first-order SDEs. In fact, the latter are plagued by the problem of interpretation, which makes ‘vanishing forces’ a fuzzy concept. Moreover, second-order equations are formulated on a geometrically neat tangent bundle.

In one dimension, the following equation with state-dependent diffusion coefficient and vanishing environmental forces is frequently encountered [5, 19, 20]:

\[ \ddot{x}_t = -\gamma(x_t) \dot{x}_t + \sqrt{2}\sigma(e(x_t)) \zeta_t, \]  

(1)

where \( \sqrt{2} \zeta_t \) is Gaussian white noise, \( \sigma(x) = \sigma e(x)^2 \) is a state-dependent diffusion coefficient and \( \gamma(x) \) is state-dependent viscosity. We will later refer to this theory as ‘state-dependent viscosity’. Assuming the state-dependent Einstein relation \( \gamma(x) = \beta \sigma(x) \), the theory leads to a Maxwellian steady state \( \pi^* (v|\cdot) = P^*(x,v) \propto \exp(-\beta v^2/2) \), with uniform spatial density. Experimental evidence by Lançon [21], Volpe [22, 23] and co-workers seems to indicate that equation (1) describes the motion of an overdamped colloidal particle when the bath has uniform temperature, while displaying a nonuniform density due to hydrodynamic interactions with confining walls.

Less studied [24, sections 1.3.2.1, 4.1.2.2] [25]–[27] is the following equation:

\[ \ddot{x}_t = -\gamma \dot{x}_t + \sqrt{2}\sigma(e(x_t)) \zeta_t. \]  

(2)

It was originally proposed by Van Kampen [28] to realize a theory settling Landauer’s physical intuitions [29] about systems with varying temperature, and therefore we will later refer to this theory as ‘state-dependent temperature’. In fact, it is obtained from the state-independent Langevin equation by simply replacing \( \beta \) with \( \beta(x) = e(x)^2 \gamma/\sigma \). It differs from equation (1) in one respect: the viscosity \( \gamma \) is not itself state dependent. It leads to a nonuniform steady distribution of particles \( p^*(x) \propto \beta(x) \), in accordance with the phenomenon of thermophoresis which drives particles along the temperature gradient, from hotter to colder zones.

We advance yet another equation:

\[ \ddot{x}_t = -\gamma \dot{x}_t + \left[ \partial_x \log e(x_t) \right] \dot{x}_t^2 + \sqrt{2}\sigma(e(x_t)) \zeta_t. \]  

(3)

As the latter theory, it involves a state-independent viscosity. However, an additional quadratic term in the velocity appears—somewhat analogous to Rayleigh’s drag force. The theory accommodates Einstein’s relation and a Maxwellian steady conditional probability \( \pi^*(v|x) = P^*(x,v)/P^*(x) \) for velocities at a given position, while the spatial density is generally not uniform, dispensing with the equal-a-priori postulate. In [18] we argue that this might be a valid alternative theory for systems with nonuniform temperature, even more adherent to Landauer’s thermodynamic picture. In the present paper we discuss the geometrical properties of the theory and of its n-dimensional generalization.

In the overdamping limit \( \gamma, \sigma \rightarrow \infty \) at fixed \( \beta \), the viscous damping term is responsible for the reduction of Langevin-type equations to first-order equations over state space, where the inertia \( \ddot{x} \) of the particle can be neglected. Notoriously, the latter are subject to antique [30, 31] and ongoing [5], [20]–[23], [32]–[34] debates on the ‘correct’ interpretation of the stochastic differential: Itô, Fisk–Stratonovich, Hägg–Klimontovich, or other? Because different stochastic differentials are mapped one into another by adding effective drift terms, the problem of interpretation is equivalent to the choice of drift in a given
fixed convention; throughout this paper we choose to work with Itô calculus. We talk of a pure SDE when the equation has no drift in some convention. Remarkably, the generally covariant first-order SDE stemming from our theory is not pure, as it requires an effective drift in all conventions, except for the one-dimensional case, where it reduces to the Stratonovich differential. It coincides with the mathematicians’ definition of Brownian motion on a manifold.

One remarkable property of our construction is that the steady state of the generally covariant reduced Fokker–Planck equation always satisfies detailed balance, with vanishing steady currents, while other approaches might lead to nonequilibrium steady states even in the absence of environmental forces, with steady currents driven by inhomogeneity of the diffusion coefficient. In particular, this is always the case for all pure stochastic differentials.

1.3. Plan and notation

The paper is organized as follows. In section 2 we discuss the statistical meaning of gauge invariance under rotations and derive the generally covariant Langevin equation, the corresponding Kramers generator and its steady state. Section 3 is devoted to the Fokker–Planck equation and to the corresponding first-order SDE deriving from the overdamping limit, which is carried out in appendix B (for a slightly more general theory). We also define covariant thermodynamic forces, currents, entropy and entropy production. In section 5 we compare our results with other theories of state-dependent diffusion and provide comparative tables in appendix D. Finally we draw conclusions in section 6.

In the following, $x, x'$ are local coordinates on an $n$-dimensional manifold $M$, which is assumed to be compact without boundary to avoid problems with conservation of probability; $x_t, x'_t$ are coordinates of a stochastic trajectory; explicit dependences on coordinates and time are often omitted; indices $a, b, c$ and $i, j, k$ label respectively coordinate and Euclidean directions; the Einstein convention on index contraction is used; spatial derivatives are abbreviated with $\partial_a = \partial/\partial x^a$; tensor indices are raised and lowered with the metric $X_a = g_{ab} X^b$, internal indices with the Euclidean metric $X_i = \delta_{ij} X^j$. Fields and probabilities are assumed to be smooth, so that there is no nonequilibrium driving due to topological defects such as sinks and sources.

2. Brownian motion

2.1. Gauge invariance under rotations

We work with $n$ independent sources $\zeta^i_t$ of Brownian white noise. They are the formal time derivatives of $n$ independent Wiener processes $(W^i_t)_t$. Wiener increments $dW^i_t = \zeta^i_t dt$ have null average, are uncorrelated at different times and in different internal directions,

$$\langle dW^i_t \rangle = 0, \quad \langle dW^i_t dW^j_{t'} \rangle = \delta^{ij} \delta(t-t') dt,$$

and are normally distributed with covariance $\delta^{ij} dt$:

$$\text{prob} (dW^1_t, \ldots, dW^n_t) \propto \exp \left( -\frac{\delta^{ij} dW^i_t dW^j_t}{2 dt} \right).$$

$$\text{prob} (dW^1_t, \ldots, dW^n_t) \propto \exp \left( -\frac{\delta^{ij} dW^i_t dW^j_t}{2 dt} \right).$$

$$\text{prob} (dW^1_t, \ldots, dW^n_t) \propto \exp \left( -\frac{\delta^{ij} dW^i_t dW^j_t}{2 dt} \right).$$
Generally covariant state-dependent diffusion

All three properties are invariant under rotations of the Wiener increments, $dW^i \rightarrow R^i_j dW^j$, where $R_{ij}$ is a special orthogonal matrix,

$$R^i_j R^j_k \delta_{ik} = \delta_{ij},$$

with unit determinant $R = \det(R_{ij})_{i,j} = 1$. The reader should note the collocation of indices to avoid confusion between a matrix and its transpose/inverse. For example, $R^i_j R^j_k = \delta^i_k$ while $R^i_j R^j_k = (R^2)^i_k$. Invariance under global rotations is obvious for equation (5) and is easily proved for equations (4) from the linearity of the average.

The gauge principle suggests consideration of the analogous local symmetry. The intuitive picture is as follows. Wiener increments are attached to each point of state space. As a Brownian particle passes by $x_t$ at time $t$ it receives a kick. While single trajectories do change if we rotate the Wiener increment at that point by a rotation matrix $R^i_j(x_t)$, the statistics of kicks remains unchanged. In fact, linear combinations of independent multivariate normal variables have multivariate normal distribution, with average

$$\langle R^i_j(x_t) dW^i_t \rangle = \langle R^i_j(x_t) \rangle \langle dW^i_t \rangle = 0,$$

and covariance matrix

$$\langle R^i_k(x_t) dW^k_t R^j_l(x_t) dW^l_t \rangle = \langle R^i_k(x_t) R^j_l(x_t) \rangle \langle dW^k_t dW^l_t \rangle = \delta^{ij} \delta(t - t') dt.$$ 

In both cases we used the properties of Itô calculus, which we conventionally work with. More precisely, any product $h(x_t) dW^i_t$ of a state function by a stochastic increment requires an interpretation about when the state function is evaluated within the time interval. The interpretation stems from the underlying discretization procedure. The trajectory is sampled at discrete times $(t_n)_n$, when it occupies states $(x_n)_n$. At intermediate times $t \in [t_n, t_{n+1})$ one approximates $h$ according to $h(x_t) = h(x_n + \alpha(x_{n+1} - x_n))$, where: $\alpha = 0$ (Itô) has a causalistic, non-anticipatory character; $\alpha = 1/2$ (Fisk–Stratonovich) is middle-point and preserves the rules of calculus; $\alpha = 1$ (Hänggi–Klimontovich) has a finalistic character. We emphasize that the choice of Itô calculus is purely conventional and does not affect the generality of the results. More importantly, this choice is not related to the dispute that we hinted at in section 1.

The advantage of Itô calculus is that the statistical properties of Wiener increments are independent of the state $x_t$ where they are nested, hence we are allowed to factor as we did in equations (7) and (8). What these equations then tell us is that the rotated Wiener increments have the same statistical properties as the starting ones: they are statistically indistinguishable. This is a gauge transformation in that it acts on the internal properties of the system—we do not measure the stochastic increments themselves—and it depends differentiably on the state.

Because Wiener increments enjoy this symmetry, we expect that any theory whose statistical properties are dictated by the statistics of Wiener increments also enjoys this symmetry. As we will point out, just as theories that are not coordinate covariant depend on the particular coordinates chosen, theories which do not satisfy gauge symmetry have very different properties according to how the symmetry is made to break. Gauge invariance of overdamped Brownian motion has also been recently discussed in [37].

doi:10.1088/1742-5468/2013/07/P07005
2.2. Generally covariant Langevin equation

While the manifold of states $M$ needs not be physical space, for the sake of visualization the Langevin equation

$$\dot{x}^a = v^a, \quad \dot{v}^a = -\gamma v^a + F^a(x,v) + \sqrt{2}\sigma e^a_i(x)\zeta^i,$$

is thought to describe the motion of a Brownian particle with mass $m$. The trajectory is driven by an effective drift $F^a$, is damped by a viscous force $-\gamma \dot{x}^a$, and receives stochastic kicks from the $\zeta^i$. Notice that we assumed that the number of degrees of freedom (manifold dimensions) and the number of noise sources is the same. In section 4.3 we show that when there are more sources of noise than space dimensions the dynamics of the system is not affected, while when there are less problems arise.

The frame $e^a_i$ mixes $n$ independent Wiener increments to yield $n$ coordinate increments. It determines the intensity and direction of random kicks, in a way that depends on the state visited by the trajectory. Correlations are characterized by the diffusion matrix

$$g^{ab}(x) = e^a_i(x)e^b_j(x)\delta^{ij}. \quad (10)$$

We assume that $(e^a_i)_i,a$ is a square invertible matrix—that is, the $n$ frame vectors $e_i = (e^a_i)_a$ are linearly independent. Its inverse $(e^i_a)_i,a$ is called the coframe. It follows that the diffusion matrix is positive definite, with inverse $(g_{ab})_{a,b}$ and inverse determinant $g = \det(g_{ab})_{a,b} > 0$. Equation (10) is defined in a local coordinate system and with respect to a local smooth frame covering an opportune neighborhood of the manifold. A coordinate systems defines (holonomic) coordinate tangent vectors that in general are not orthogonal, while frame vectors are orthogonal (but generally anholonomic). The base manifold might not admit a coordinate system with smooth tangent vectors all over the manifold (i.e. a global section of the tangent or of the frame bundles), so that from time to time it is necessary to move to a new neighborhood by changing coordinates and smooth frames. Covariance guarantees that where the neighborhoods overlap the equations are consistent. See section 4.2 for more considerations and an example.

We now perform an invertible coordinate transformation $x \rightarrow x'(x)$ and, by the principle of covariance, we assume that the trajectory $x'_i = x'(x_i)$ obeys an equation in the same form as equation (9):

$$\dot{x}'^a = v'^a, \quad \dot{v}'^a = -\gamma v'^a + F'^a + \sqrt{2}\sigma e'^a_i(x)\zeta^i. \quad (11)$$

In theory, we should resort to the rules of Itô calculus to evaluate differentials; in practice, for this kind of equation stochastic and standard calculus coincide, and there is no issue concerning the interpretation of the stochastic differential. From the first transformed equation $\dot{x}'^a = v'^a \partial_a x'^a$ we read off the transformation law for the velocity $v'^a = v^a \partial_a x'^a$, which behaves like a vector sitting in the tangent space at $x$. We then take the stochastic differential of $v'^a$, using the Itô/Leibniz rule

$$dv'^a = v^a \, d\partial_a x'^a + dv^a \, \partial_a x'^a = (\partial_a \partial_{b} x'^a v^{a} v^{b} + \partial_{a} x'^a F^{a}) \, dt + \sqrt{2} \partial_{a} x'^a e'^a_i \zeta^i \, dt, \quad (12)$$

which leads to

$$e'^a_i = \frac{\partial x'^a}{\partial x'^a} e^a_i, \quad F'^a = \frac{\partial x'^a}{\partial x'^a} F^{a} + v^a v^{b} \frac{\partial^2 x'^a}{\partial x'^a \partial x'^b}. \quad (13)$$

doi:10.1088/1742-5468/2013/07/P07005
The frame’s components, at fixed \( i \), transform like vectors, while the inverse diffusion matrix is twice contravariant, \( g_{ab} = g_{d'a'b}x^d\partial_a x^b \). Therefore, the latter bears the properties of a metric, with invariant volume element \( \sqrt{g} dx \). Because \( g_{ab} e^a_i e^b_j = \delta_{ij} \), the frame vectors \( e_i \) form a basis of orthonormal vectors for the tangent space at \( x \), whence the interpretation as orthonormal reference frame, point-by-point over the manifold. In general, such reference frames are anholonomic (noncoordinate), meaning that the matrix \( e^a_i \) is not the Jacobian of a coordinate transformation that makes \( e^a_i = \delta^a_i \) and \( g_{ab} \) the flat Euclidean metric, but at one point on the manifold at a time. The frame makes its appearance in general relativity, where it is known as vielbein (counting dimensions in German, viel = zwei, drei, vier, etc). For example, the vierbein maps the local flat 4D flat Minkowski metric—three sticks and a clock, following Einstein’s physical intuition—into a semi-Riemannian metric on the spacetime manifold.

As is well known, the inhomogeneity developed by the drift term in equation (13) qualifies it as

\[
F_1^a = -\Gamma^a_{bc} v^b v^c, \tag{14}
\]

where \( \Gamma^a_{bc} \) are the Christoffel symbols of the Levi-Civita connection, i.e. the unique torsionless connection compatible with metric \( g_{ab} \),

\[
2\Gamma^a_{bc} = g^{ad} (\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}). \tag{15}
\]

However, this is not the only viable choice, as

\[
F_2^a = v^c v^b \tilde{e}^b_i \partial_c e^a_i \tag{16}
\]

produces the same inhomogeneous term. As we will see later, in the overdamping limit the latter leads to Graham’s theory while the former leads to the mathematicians’ Brownian motion on a manifold.

For a sensible choice between the two, we invoke the principle of gauge invariance. We assume that the theory is invariant under internal \( SO(n) \)-transformations of the reference frames, acting on the \( i, j, k \) indices

\[
e^a_i (x) \rightarrow \tilde{e}^a_i (x) = e^a_j (x) R^j_i (x), \quad e^a_i (x) \rightarrow \tilde{e}^a_i (x) = e^a_i (x) R^i_j (x), \tag{17}
\]

where the transformation law for the coframe follows from \( \tilde{e}^a_i \tilde{e}^j_b = \delta^a_b \). Notice that the metric is invariant under local coordinate transformations, \( \tilde{g}_{ab} = g_{ab} \), and so are the Christoffel coefficients. In general relativity, gauge transformations have the physical meaning of rotations of the observer’s orthogonal reference frame. As we argued in the dedicated paragraph, \( \zeta^i_t \) and \( \tilde{\zeta}^i_t = R^i_j (x) \zeta^j_t \) are statistically indistinguishable, hence the stochastic term is invariant under \( e^a_i \zeta^i_t \rightarrow \tilde{e}^a_i \tilde{\zeta}^i_t = e^a_i \tilde{\zeta}^i_t \).

Given that the stochastic term is statistically preserved by gauge transformations, it remains to check that the deterministic term is. As a rule of thumb, it should be expressible only in terms of the metric and of its derivatives. This is the case for \( F_1 \), while it is not for \( F_2 \). In fact, after a gauge transformation,

\[
\delta F_1^a = 0, \quad \delta F_2^a = e^a_k e^c_b v^b v^c \partial_\zeta R^k_i. \tag{18}
\]

Therefore we choose \( F_1 \), obtaining

\[
\dot{v}^a + \Gamma^a_{bc} v^b v^c = -\gamma v^a + \sqrt{2} \sigma e^a_i \zeta^i. \tag{19}
\]

doi:10.1088/1742-5468/2013/07/P07005
In the deterministic case $\gamma = \sigma = 0$ one recognizes the geodesic equation. Equation (19) is our proposal for a Langevin equation with state-dependent diffusion matrix. On a flat manifold, in coordinates where the metric is Euclidean and the Christoffel coefficients vanish, it returns the ordinary Langevin equation.

At the time of submitting, the author became aware that equation (19) could already be found in a work by Kleinert and Shabanov [38], who discussed its generalization to connections with torsion. Very recently the work of Castro-Villarreal and co-workers [39] has also been pointed out—they derived it using an extrinsic, rather than intrinsic, approach.

2.3. Kramers’ equation and its steady state

While the $n$-dimensional process $(x_t)_t$ is not a Markovian diffusive process, the $2n$-dimensional process $z_t = (x_t, v_t)$ on the tangent bundle $TM$ is. By a standard procedure, the Kramers equation is obtained by writing down the diffusion equation for the probability density $P_t(z)$ and expressing it in terms of $(x,v)$:

$$\dot{P} = -v^a \frac{\partial P}{\partial x^a} + \frac{\partial}{\partial v^a} [(\Gamma^a_{bc} v^b v^c + \gamma v^a) P] + \sigma g^{ab} \frac{\partial^2 P}{\partial v^a \partial v^b}. \quad (20)$$

The steady state is given by

$$P^*(x,v) = \frac{g(x)}{N} \exp \left( -\frac{\gamma}{2\sigma} g_{ab}(x)v^a v^b \right). \quad (21)$$

This can be checked by direct substitution—the calculation entails an interesting balancing of Christoffel terms $\Gamma_{abc} v^a v^b v^c$ and $2\Gamma_{ab} v^a$ from $v$ to $x$ derivatives.

Since the metric has physical dimensions $[x]^{-2}$, the viscosity and the diffusion constant have dimensions respectively $[\gamma] = [t]^{-1}$ and $[\sigma] = [t]^{-3}$. We define an energy,

$$H(x,v) = \frac{m \lambda_c^2}{2} g_{ab}(x)v^a v^b, \quad (22)$$

where we introduce the Compton wavelength of the Brownian particle $\lambda_c = \frac{2\pi \hbar}{mc}$ for the sake of dimensional consistency. In appendix A it is shown that the average energy $E_t = \langle H(x_t, v_t) \rangle$ is a Lyapunov functional, exponentially decreasing to the steady value $n\beta^{-1}/2$, where

$$\gamma/\sigma = \beta (2\pi \hbar)^2/mc^2. \quad (23)$$

The identification of $\beta$ with an inverse invariant temperature is suggested by the equipartition theorem at the Gibbs steady state $P^* \propto \exp(-\beta H)$, yielding the Einstein relation. Upon coordinate transformations, the energy is a scalar, the temperature is invariant and the probability density transforms like a squared volume density (like $g$), consistently with the fact that $P_t(x,v) \, dx \wedge dv$ ought to be an invariant volume element to preserve normalization.

Finally, we calculate the steady spatial density by integrating over velocities:

$$p^*(x) = \int_{-\infty}^{+\infty} P^*(x,v) \, dv = \frac{\sqrt{g(x)}}{\int_M \sqrt{g(x')} \, dx'}, \quad (24)$$

doi:10.1088/1742-5468/2013/07/P07005
where \( N = (mc^2/2\pi\hbar^2)^{n/2} \int_M \sqrt{g(x')} \, dx' \). Brownian particles tend to occupy space according to the volume element. More probable states occur where noise is less intense. In fact, if we modulate the intensity of noise without modifying the angles between frame vectors, by performing a rescaling \( \epsilon_i \rightarrow \lambda_i \epsilon_i \) (no index contraction), the determinant of the metric scales like \( g \rightarrow g/\prod_i \lambda_i \). This is physically consistent, as more intense noise kicks particles further away, making permanence at a state briefer. Modifying the angles, \( g \) increases when the frame vectors are made decreasingly linearly independent, shooting to infinity when \( (\epsilon_a^b)_{i,a} \) becomes degenerate.

The conditional probability density \( \pi^*(v|x) = P^*(x,v)/P^*(x) \), describing how velocities in the tangent space \( T_x M \) distribute at a fixed \( x \), is Maxwellian. However, because \( x \) and \( v \) terms in \( P^* \) do not factor, the velocity probability density \( \Pi^*(v) = \int_M P^* \, dx \) will generally not be Gaussian. This is a distinctive signature of the theory.

3. Overdamped dynamics

3.1. Fokker–Planck equation

The overdamping (Smoluchovski) limit of the Langevin equation is achieved by sending \( \gamma, \sigma \rightarrow \infty \) at fixed \( \beta \), in which case the damping effect of the viscous force is so strong that the inertial term \( \dot{v}^a \) drops. The limit must be taken with great care; details of a rigorous derivation are reported in appendix B, where the Kramers equation is shown to reduce to the following Fokker–Planck equation:

\[
\dot{\rho}_t = \partial_a \left[ g^{ab} \left( \partial_b \rho_t - \rho_t \partial_b \log \sqrt{g} \right) \right]
\]  

(25)

for the spatial probability density

\[
\rho_t(x) = \int dv \, P_t(x,v).
\]  

(26)

Introducing the covariant derivative \( \nabla_a \) associated with the Levi-Civita connection, defined by \( \nabla_a X^b = \partial_a X^b + \Gamma^b_{ac} X^c \), equation (25) can be recast in manifestly covariant form

\[
\dot{\rho} = \Delta_M \rho - \nabla_a (f^a \rho),
\]  

(27)

where \( \rho_t = \rho_t/\sqrt{g} \) is a scalar and \( \Delta_M = \nabla^a \nabla_a \) is the Laplace–Beltrami operator. For the sake of generality we have momentarily accounted for an environmental force vector \( f^a \). In deriving equation (27), we used the fact that the covariant derivative of a scalar coincides with its partial derivative and that the covariant gradient is given by \( \nabla_a X^a = \sqrt{g}^{-1} \partial_a (\sqrt{g} X^a) \). When \( f^a = 0 \), equation (27) returns the well known definition of Brownian motion on a manifold in the specialized mathematics literature (up to a factor 1/2) [7, chapter 3] [8, section 4.2.2]; it was also considered by Batrouni et al [40] with particular interest in the stochastic quantization of lattice gauge theories, and by Castro Villarreal [41], who worked out the small-time/low-curvature dynamics.

3.2. Smoluchovski equation

The SDE that corresponds to equation (25) is given by

\[
\dot{x}^a = f^a + f^a_{\text{eff}} + \sqrt{2 \epsilon_i^a} \zeta^i,
\]  

(28)

doi:10.1088/1742-5468/2013/07/P07005
Generally covariant state-dependent diffusion

where again we have included the environmental force vector, and the effective drift reads

\[ f^{a}_{\text{eff}} = -g^{bc} T^{a}_{bc} = \partial_b g^{ab} + g^{ab} \partial_b \log \sqrt{g}. \]  

Equation (28) can also be derived from first principles in much the same way as the generally covariant Langevin equation, taking into account that by the rules of Itô calculus the effective drift transforms as

\[ f^{a'}_{\text{eff}} = \partial_a x^{a'} (f^{a}_{\text{eff}} + g^{ab} \partial_a \partial_b x^{a'}). \]

In short, one performs a Taylor expansion of \( x^{a'}(x^{t}) \) to second order in \( dx^{t} \), plugs equation (28) in and retains terms of order \( dt \), keeping in mind that squared Wiener increments \( (dW^{t})^2 \) equal \( dt \) with certainty.

Interestingly, when \( f^{a} = 0 \), for no value of \( \alpha \) does equation (28) coincide with a pure SDE

\[ \dot{x}^{t} = \sqrt{2} e^{a}_{i} (x^{t} + \alpha \, dx^{t}) \zeta^{t}_{i}, \]  

as there always appears an effective drift term, but in one dimension, where it reduces to the Stratonovich differential. We notice in passing that the Itô differential leads to a gauge-invariant but not coordinate-covariant SDE, while the converse occurs for the Stratonovich convention. In all cases, upon either a coordinate or a gauge transformation, such differentials usually gain an additional effective drift. More interestingly, pure SDEs in equation (31) usually lead to nonequilibrium steady states, as we briefly argue in section 5.

3.3. Covariant entropy and entropy production

Entropy and entropy production are central to macroscopic thermostatics and thermodynamics. We give covariant analogues in first-order theory, while a complete thermodynamic treatment is reserved for a future publication. Covariant equilibrium thermostatics has also been recently discussed in [37].

From an information-theoretic point of view, the uniform distribution maximizes the Gibbs–Shannon entropy, a measure of ignorance with respect to a reference prior, given by the uniform distribution itself. The Gibbs–Shannon entropy notoriously develops an inhomogeneous term when one keeps the uniform prior fixed through a coordinate change, \( \delta S = - \int \rho \log \det \, \partial x' \, dx \). This can be avoided by transforming the prior as well, which will therefore not be uniform in the new coordinates.

Thus, it does not make much sense to pick the uniform prior in some preferred coordinates in the first instance. Our theory comes with its own preferred prior, namely the volume element \( \sqrt{g} \). A natural covariant candidate as a measure of excess information is then

\[ D_{\sqrt{g}} = \int \, dx \, \sqrt{g} \rho \log \rho, \]  

where \( \rho = p/\sqrt{g} \) is a scalar. The above expression is relative entropy with respect to \( \sqrt{g} \). Though not uniform (a matter of concern for some authors [42]), \( p^{*} \) is analogous to the microcanonical ensemble, with equal volumes in state space having equal probabilities. There is an obvious circularity in this argument, since the choice of \( \sqrt{g} \) as the preferred volume element is subjective, as is equiprobability. In fact, \( \sqrt{g} \) bears a resemblance to
Jeffrey’s prior in Bayesian inference, whose objective versus subjective character is still debated [43].

The author recently argued that subjective priors have full citizenship in NESM [17]. A change of prior does modify the thermostatics of the system but not its thermodynamics, which is fully encoded in the concept of entropy production, which we now introduce. Let us deal with external forces for sake of generality. We introduce covariant currents and affinities,

\[ J^a = f^a \rho - \nabla^a \rho, \quad A_a = f_a - \nabla_b \log \rho, \quad (33) \]

and the entropy production

\[ \sigma = \int \sqrt{g} J^a A_a = \int p(\text{d}x) g_{ab} A^a A^b = \langle A^2 \rangle. \quad (34) \]

The equalities on the rhs show that the entropy production is non-negative. It vanishes at the steady state when the condition of detailed balance holds, that is when there exists a potential \( \phi \) such that \( f_a = -\nabla_a \phi \), with steady state \( \rho = \sqrt{g} \exp(-\phi) \). Notice that our free theory, with \( f_a = 0 \), satisfies detailed balance. A covariant necessary condition for detailed balance is

\[ \nabla_a f_b - \nabla_b f_a = \partial_a f_b - \partial_b f_a = 0. \quad (35) \]

Even if this condition is satisfied though, the force might not be globally a gradient, igniting topological steady currents. Notice that the covariant condition is equivalent to the noncovariant condition (second passage) by virtue of the torsionlessness of the connection. We are tempted to speculate that torsion might therefore be a nonequilibrium drive, but we leave this to future inquiry.

The entropy production can be expressed in terms of the time derivative of the relative entropy, which quantifies the internal entropy production, and an additional environmental term, \( \sigma = -D\sqrt{g} + (A^a f_a) \), where

\[ (A^a f_a) = \int p(\text{d}x) f^a \left( f_a + \partial_a \log \sqrt{g} - \partial_a \log p \right). \quad (36) \]

The second term within parentheses can be shifted from the environmental to the internal term, returning the standard Gibbs–Shannon entropy and the standard environmental term usually considered in the literature [3]. The covariant splitting of the entropy production is one of many physically meaningful splittings into internal and environmental terms, all of which can be mapped into each other by a change of reference prior. This freedom is an additional fundamental symmetry of nonequilibrium thermodynamics, deeply mingled with coordinate covariance.

4. Geometric insights

4.1. Gauge connection

The promotion of global symmetries to local ones by virtue of the gauge principle usually entails the appearance of a new field, which compensates the transformation law and introduces a new interaction. The gauge field for \( SO(n) \) transformations is the affine connection \( \omega_{aij} \), \(^1\) which generates an infinitesimal rotation as one displaces vectors along a

\(^1\) Sometimes called the spin connection in the Cartan approach to general relativity [35, 36].
given direction—namely \((\omega_{aij})_{i,j}\) is an element in the Lie algebra \(\mathfrak{so}(n)\) of skew-symmetric \(n \times n\) matrices. Here we show how our theory can be expressed nicely in terms of the affine connection and of the closely related anholonomicity vector. This is the only truly geometrical part of the paper; the reader might safely skip it. Nevertheless, we maintain the general inaccurate tone depicting geometry as index gymnastics.

Since fields are tensors, we should hereby resort to Stratonovich calculus, which preserves the rules of differential calculus. equation (28), with \(f^a = 0\), now reads

\[
\dot{x}^a = -\delta^{ij} e^b_i \partial_b e^a_j - g^{bc} \Gamma^e_{bc} + \sqrt{2} e^a_i \circ \zeta^i = -\omega^a + \sqrt{2} e^a_i \circ \zeta^i,
\]

where \(h(x_t) \circ \zeta^i_t = h(x_t + dx_t/2)\zeta^i_t\) is the Stratonovich differential. In the first passage we have added a drift term to switch convention. After some manipulation one obtains the second equation, where

\[
\omega_a = e^b_i (\partial_a e^j_b - \partial_b e^j_a)
\]

is called the \textit{anholonomicity (co)vector}. After a gauge transformation the anholonomicity vector gains an inhomogeneous term which on average balances exactly the one developed by the Stratonovich differential.

Notice that the anholonomicity vector has the typical transformation law \(G^{-1} \partial_a G\) of gauge fields, where \(G\) is a group element. Therefore, the anholonomicity vector is the gauge field of generally covariant stochastic thermodynamics. It will be present in all expressions where rotation invariance requires adjustment. In a way, the anholonomicity vector introduces an effective interaction between the Brownian particle and the environment which bends its motion:

\[
\delta \langle \omega_a \rangle = - \langle e^b_i e^j_i R^a_i \partial_e R^b_e \rangle = \delta \langle \sqrt{2} e^a_i (x_t) \circ \zeta^i_t \rangle.
\]

Let us see how the theory is expressed in terms of the affine connection. The latter

\[
\nabla_a e^i_b = \partial_a e^i_b - \Gamma^i_{ab} e^j_c + \omega^i_{a j} e^j_b = 0.
\]

Antisymmetrizing, and recalling that the Levi-Civita connection is torsionless, \(\Gamma^c_{[a,b]} = 0\), we can express the anholonomicity vector as

\[
\omega_a = \omega^i_{kj} a^k_i b^j,
\]

where we have used \(\omega^i_{a k} = \delta^{ij} \omega^i_{a j} = 0\). Finally, we can write equation (37) in a particularly simple form in terms of the affine connection and of the velocity \(v^i_t = e^i_a(x_t) \circ \dot{x}^a_t\):

\[
v^i + e^a_i \omega^i_{aj} = \sqrt{2} \zeta^i.
\]

Let us briefly discuss the geometric significance of the anholonomicity vector. It measures to what extent the orthogonal basis fails to be a coordinate basis. If there exists a collection of potentials \(u^i\) such that \(e^i_a = \partial_a u^i\), then \(\omega^a\) vanishes. Performing coordinate transformation \(x^a \rightarrow u^i(x)\), the transformed metric is easily seen to be Euclidean. In this case, the manifold is flat, and it has vanishing Riemann curvature. In general, the Riemann curvature of a manifold does not vanish, so that for such systems it is impossible.
to find coordinates where the anholonomicity vector vanishes. We skip the case where the manifold is only locally flat. We only observe that (local) flatness has no thermodynamic relevance to our theory, as far as we can see.

4.2. Nonparallelizable manifolds: the sphere

All equations above are written in a local chart of coordinates and with respect to a smooth field of frames—that is, a local section of the so-called frame bundle. Not all manifolds admit a global chart of coordinates with a globally smooth set of coordinate tangent vectors $\partial_a$. In other words, nonparallelizable manifolds do not admit a global section of both the tangent and the frame bundles. Accordingly, in such coordinates the metric might become degenerate or singular at some points, but such pathologies are artifacts of the coordinate system employed. This implies that from time to time it is necessary to switch coordinate system and, in doing so, push forward the frame vectors so as to be able to cover the full manifold with a nondegenerate frame. Covariance guarantees that the dynamics remains consistently the same where the two charts intersect. Defining the dynamics on an atlas of local charts that covers the manifold uniquely defines the theory.

For example, the Smoluchowski equations for a Brownian particle on a sphere with no external forces in spherical coordinates read (in Stratonovich convention)

$$\dot{\theta} = \frac{\cos \theta}{\sin \theta} + \sqrt{2} \zeta^1,$$
$$\dot{\phi} = \frac{\sqrt{2}}{\sin \theta} \circ \zeta^2 \tag{43}$$

where $\phi \in [0, 2\pi)$ is the azimuthal angle and $\theta \in (0, \pi)$ the polar angle ($\theta = 0, \pi$ being the poles). From these equations we read off the entries of the frame vectors $e_1 = (1, 0)$ and $e_2 = (0, \sin^{-1}\theta)$. The metric in spherical coordinates reads $d\theta^2 + \sin^2\theta d\phi^2$. Notice that at the poles the metric is degenerate, $e_1$ is not continuous and $e_2$ is singular, complying with the hairy-ball theorem which states that there is no smooth nonvanishing vector field tangent to a sphere. Correspondingly, the equations become singular. This is due to the fact that the poles $(0, \phi)$ are not uniquely defined in this coordinate system; they are fictitious pathologies (unlike, say, the tip of a cone). When approaching regions near the poles one has to employ a different coordinate system, for example by rigidly rotating the polar axis. In doing so frame vectors from regions where they are well defined are rigidly transported (pushed forward) to the poles, and one can continue the description of the system there.

4.3. Redundant and defective sources of noise

In this section we briefly consider the case when there are more or fewer sources of noise than degrees of freedom. Letting $I \in [0, m]$, $a \in [0, n]$, we consider the equation

$$\dot{x}^a = f^a + e^a_I \zeta^I \tag{44}$$

Suppose there is a redundant number of sources of noise, $m > n$. We require that the $n \times m$ matrix $e^a_I$ contains $n$ independent columns, so that the diffusion matrix $g^{ab} = \delta^{IJ} e^a_i e^b_j$ is positive definite. It is well known that any positive definite symmetric matrix can be written in Gramian form $g^{ab} = \delta^{ij} e^a_i e^b_j$ in terms of $n$ independent frame vectors, with $i \in [0, n]$. In fact, because by the spectral theorem any positive symmetric matrix can
be diagonalized by an orthogonal transformation $A$ and it has positive eigenvalues $\{g_a\}_a$, then taking $\sqrt{g_a}$ times the $a$th column of $A$ gives one such frame vector. Notice that in general $g^{ab}$ cannot be diagonalized at the same time at every point by a coordinate transformation, which implies that in general the field of orthogonal matrices $A$ is not the Jacobian of a coordinate transformation over an open neighborhood. Now define

$$\tilde{\zeta}^i = e^i_a e^a_I \zeta^I.$$  \hfill (45)

It is well known that linear combinations of independent Gaussian variables are Gaussian variables, hence $\tilde{\zeta}^i$ are $n$ sources of Gaussian white noise, with equal-time correlation

$$\langle d\tilde{W}^i \, d\tilde{W}^j \, dt \rangle = e^i_a e^j_b \langle dW^a \, dW^b \rangle = (g_i g_j)^{-1/2} g^{ab} A^i_a A^j_b = \delta^{ij}. $$  \hfill (46)

Hence the process in equation (44) and

$$\dot{x}^a = f^a + \tilde{e}^a_I \tilde{\zeta}^I $$  \hfill (47)

are statistically indistinguishable, and in particular they will have the same dynamics. So from a dynamical perspective a redundant set of sources of noise is irrelevant, since it can always be coarse grained to exactly $n$ sources of noise preserving the dynamics. At the level of the thermodynamical description, instead, it is known that similar coarse graining procedures systematically underestimate the entropy production [44], as one has to distinguish independent contributions to the heat.

The case $m < n$ can be seen as a system with constraints. We do not analyze it in detail here, as it requires the whole theory to be reformulated: the metric becomes non-invertible, hence no Christoffel coefficients can be defined, and so forth. Yet, one should be able to properly define covariant dynamics, which will have many odd peculiarities. In particular we expect that while the dynamics might be ergodic (in the sense that it will finally explore the full manifold) the steady measure might not be unique.

5. Comparison of models

We consider several models of state-dependent diffusion with vanishing external forces that can be found in the literature, and compare them with (A) our model: (B) Graham’s covariant theory of nonequilibrium statistical mechanics, (C) diffusion in nonuniform temperature, (D) diffusion with state-dependent viscosity. In the overdamping limit (B) and (C) lead respectively to Stratonovich and Itô pure SDEs. For the sake of completeness, we also consider (E) the pure SDE with posticipatory character (Klimontovich convention). Finally, we include two more general theories: (F) a noncovariant theory with state-dependent viscosity and temperature, and (G) its covariant counterpart, a theory where the diffusion matrix and the metric need not necessarily coincide. The latter includes both our theory and diffusion with state-dependent viscosity as special cases. In appendices B and C the overdamping limit for the latter two theories are performed. All others follow as special cases. We summarize all these theories’ properties in tables D.1 and D.2 and through the interdependence diagram in appendix D. Notice that (G) precedes (D) and (F) in the discussion. In this section we will set $(2\pi\hbar)^2/mc^2 = 1$. 

doi:10.1088/1742-5468/2013/07/P07005
5.1. (B) Second-order Graham theory

Suppose that $F_1$ is replaced by $F_2$ in our theory. By taking the overdamping limit, along the same lines exposed in appendix B, one would obtain Graham’s first-order theory [10]. Therefore, we can already conclude that the theory is coordinate covariant but it is not gauge invariant. The derivation leads to the following Stratonovich pure differential equation:

$$\dot{x}^a = \sqrt{2} e^a_i \circ \zeta^i = \delta^{ij} e^b_j \partial_b e^a_i + \sqrt{2} e^a_i \zeta^i.$$  \hfill (48)

The corresponding manifestly covariant FP equation is

$$\dot{\rho} = \nabla^a (\nabla_a \rho - \omega^a \rho),$$  \hfill (49)

recalling that $\omega^a$ is the anholonomicity vector, defined in equation (38). As expected, the theory is not gauge invariant since the anholonomicity vector is not. Hence, knowledge of the covariance matrix $g^{ab}$ is not sufficient for determining the theory. All statistical properties depend on the choice of frames $e^a_i$, including the equilibrium/nonequilibrium character. If for some choice of frame the theory satisfies detailed balance, $\omega_a = -\nabla_a \phi$, because the inhomogeneous term in equation (39) is not generally a gradient, the gauge-transformed theory might not satisfy detailed balance. As an example, consider a two-dimensional theory with $e^a_i = \delta^a_i$ and $\omega_a = 0$ in some gauge and coordinates. In two dimensions $R^i_j(x) = R^i_j(\theta(x))$ is a rotation matrix by an angle $\theta(x)$. To first order in $\theta$ the transformed anholonomicity vector has a contribution $\delta(\omega_1, \omega_2) = (-\partial_2 \theta, \partial_1 \theta)$, which is certainly not a gradient.

This problem with Graham’s theory was also recognized in [45].

5.2. (C) State-dependent temperature

To our knowledge, theories with state-dependent diffusion coefficient and uniform viscosity have been considered in the literature only in one dimension [25]–[28], where they have been used as models of systems with varying temperature. Let us generalize to $n$ dimensions:

$$\ddot{x}_i^a = -\gamma \dot{x}_i^a + \sqrt{2} \sigma e_i^a(x_i) \zeta_i^a.$$  \hfill (50)

In this quite general case temperature might not be isotropic, besides being inhomogeneous: it depends on the direction of motion of the Brownian particle. Given the considerations developed in this paper, the theory is gauge invariant but it is not coordinate covariant. The corresponding Kramers generator

$$\dot{P} = -v^a \frac{\partial P}{\partial x^a} + \frac{\partial}{\partial v^a} (\gamma v^a P) + \sigma g^{ab} \frac{\partial^2 P}{\partial v^a \partial v^b}$$  \hfill (51)

does not admit a locally Maxwellian steady state, as one can appreciate by plugging $\exp[-h_{ab}(x)v^a v^b / 2 + k(x)]$ as an ansatz. First- and third-order terms in the velocities can only vanish if $k = 0$ and $\partial_a g_{bc} = 0$. The exact form of the steady state of equation (51) is not easily computable to our knowledge.

doi:10.1088/1742-5468/2013/07/P07005
As for Graham’s theory, in general the steady state does not satisfy detailed balance. The nontrivial overdamping limit is performed in detail in appendix C, leading to the Fokker–Planck equation in Itô form:

$$\dot{p} = -\partial_a j^a = \partial_a \partial_b \left( g^{ab} p \right).$$

(52)

The steady current is $$j^a_* = -p^a (g_{ad} \partial_c g^{cd} + \partial_a \log p^*).$$ For it to vanish, hence for the steady state to be in equilibrium, it is necessary that $$g_{ad} \partial_c g^{cd}$$ be an exact differential. However, this is not generally the case. Detailed balance does hold when temperature is isotropic, by which we mean that there exist coordinates such that

$$g_{ab}(x) \propto T(x)^{-1} \delta_{ab}.$$  

(53)

In this case the metric is said to be conformally flat.

In a parallel publication we argue on physical grounds that our model should be considered as a valid alternative to diffusion in nonuniform temperature [18].

5.3. (G) Metric/diffusion theory

One peculiarity of our theory is that the inverse diffusion matrix coincides with the metric—noise determines the form of deterministic geodesic motion. This provides an elegant theory, but it is not strictly mandatory for a generally covariant theory.

We might introduce a new nondegenerate metric $$h_{ab},$$ with Christoffel coefficients $$\Omega^a_{bc}.$$ Let the diffusion matrix $$g^{ab}$$ be a twice covariant symmetric tensor, possibly different from the inverse metric $$h^{ab}$$. The equation

$$\ddot{x}^a + \Omega^a_{bc} \dot{x}^b \dot{x}^c = -\gamma^a_{bc} \dot{x}^b + \sqrt{2} \sigma e_i \zeta,$$

(54)

with $$g^{ab} = \delta_{ij} e^a_i e^b_j,$$ is generally covariant when $$\gamma^a_{bc}$$ is a tensor. We assume the Einstein relation $$\gamma^a_{bc} = \beta \sigma g^{ac} h_{cb}$$. The diffusion matrix and the metric govern respectively the noisy and the deterministic behavior. The viscosity coefficient is a tensorial object, once covariant and once contravariant. It is a hybrid of the diffusion matrix and of the metric, interpolating between stochastic and deterministic behavior. The Kramers equation affords the equilibrium steady state

$$P^* = h \exp(-h_{ab} v^a v^b / 2).$$

(55)

Notice that the metric, and not the diffusion matrix, determines the form of the steady state. Our theory is recovered by setting $$h_{ab} = g_{ab}$$. The overdamping limit is performed in appendix B, yielding

$$\dot{p} = \partial_a \left[ h^{ac} g_{cd} h^{db} \left( \partial_b p - p \partial_b \log h \right) \right].$$

(56)

5.4. (D) State-dependent viscosity

Theories with state-dependent diffusion and viscosity have been widely studied. See [5] for a review. The n-dimensional case, with a rigorous treatment of the overdamping limit, is discussed in [19], while [46] gives a microscopic derivation in the spirit of the Caldeira–Leggett model and [47, 48] discuss the path integral formulation.
The free model is a special case of the above tensor viscosity model when \( h_{ab} = \delta_{ab} \) and \( \Omega_{bc}^a = 0 \). This identification can only hold in one given set of coordinates, hence the theory is not covariant—although it is gauge invariant. The steady state is the usual Maxwell–Boltzmann state, displaying no dependence on the spatial variable—in a way, it is blind to the inhomogeneity of the medium. In the overdamping limit, the theory leads to the SDE
\[
\dot{x}_t^a = \partial_b g^{ab} + \sqrt{2} \epsilon^a_i \zeta_t^i.
\]
(57)
The Fokker–Planck generator is \( \dot{p} = \partial_a (g^{ab} \partial_b p) \), sometimes named after Fourier and Fick. The steady state always satisfies detailed balance. In one dimension equation (57) reduces to a pure SDE in the Klimontovich interpretation.

5.5. (E) Pure posticipatory SDE

Theories (B) and (C) led respectively to the Stratonovich and to the Itô differential in the overdamping limit. For the sake of completeness, we also consider the pure SDE in Klimontovich form:
\[
\dot{x}_t^a = \sqrt{2} \epsilon^a_i (x_t + dx_t) \zeta_t^i = 2 \delta_{ij} \epsilon^b_i \partial_b \epsilon^a_j + \sqrt{2} \epsilon^a_i (x_t) \zeta_t^i.
\]
(58)
The theory is neither covariant nor gauge invariant. Moreover, the drift term arising in the overdamping limit, as for the Stratonovich and the Itô case, is not generally a gradient, so the theory does not accommodate detailed balance as an essential feature. It is remarkable that, in higher dimensions, none of the pure stochastic differentials generally afford equilibrium steady states.

5.6. (F) State-dependent viscosity and temperature.

To complete the collection, we recall that Van Kampen considered a one-dimensional generalization of state-dependent temperature and state-dependent viscosity to the case where both are state dependent [49]. Jayannavar and Mahato [50] claimed to have derived the theory from a microscopic model of state-dependent interactions with a bath, although it appears that the state dependence of the temperature is assumed along the way, rather than being derived. Properties of the theory are summarized in the tables in appendix D.

6. Conclusions

There are two viable perspectives on the results hereby derived. On the one hand, they might be seen as an exercise on abstract symmetry principles, later to be applied to theories that already enjoy diffeomorphism invariance, such as the physics of curved spacetimes and general relativity. We point out that relativistic diffusion is a broadly studied subject, and that it has some crucial peculiarities that do not allow direct application of our results. In this respect, although our Langevin equation is similar to equation (51) in [51] (but for a term \( 2 \Gamma_{ab}^{cb} \eta_{P}^c \)), their meaning is very different as to the role of time and the relationship between the diffusion matrix and the metric. Still, covariant diffusion has an interest of its own. Graham introduced his own theory with these words: ‘Physical properties are independent of the coordinates used. Hence
Generally covariant state-dependent diffusion

they must be formulated in terms of covariant quantities’ [10]. The will to ‘covariantize’ (notions of) nonequilibrium thermodynamics was the ultimate drive for the present work as well. Recently Smerlak theorized that, along with propagation, it might be possible to ‘tailor’ the diffusion of light in suitably engineered metamaterials [52]. Our proposal and his belong to the same paradigm, that spatially dependent diffusion coefficients might act as an effective curvature of space(time).

On the other hand, our work might be seen as a doorway to state-dependent diffusion even for mesoscopic systems, both at the level of interpretation and for physical modeling. By no means do we claim that the generally covariant Langevin equation is the correct equation for state-dependent diffusion, but it should be in all situations where the statistical properties of noise do not dependent on one preferred coordinate system. This demanding hypothesis has to be addressed independently, for example when performing microscopic derivations. A class of systems where curved Brownian motion is relevant is protein transport inside and across cell membranes under the effect of membrane curvature (see the introduction of [39] and references therein). In a parallel publication we also argue that it might be relevant for Brownian motion in nonuniform temperature environments [18]. The latter conjecture is already testable, as experimental works by Lançon et al [21] and by Volpe et al [22] (after a few issues have been resolved [23]) display great control over Brownian particles under the effect of spatial dependence of the diffusion coefficient, in this case due to the presence of walls. It might be feasible to design similar experiments where spatial dependence is due to temperature gradients or both to temperature gradients and spatial inhomogeneity, and put several of the models we described above at work. The generally covariant Langevin equation could also find application in Metropolis Langevin Monte Carlo sampling methods, according to the lines of research developed in [53].

The relationship between gauge invariance and detailed balance has been a main theme in this paper. It follows from our discussion that gauge invariance is necessary (but not sufficient) to write thermodynamic theories of state-dependent diffusion that enjoy equilibrium steady states, as different choices of gauge will otherwise lead to nonequilibrium driving. This was the case for Graham’s theory. Surprisingly, instead, curvature does not affect the equilibrium character of steady states; it might be interesting to consider instead the thermodynamic effect of geometric torsion.

To conclude, we point out that most thermodynamic considerations in this work were carried out at the macroscopic level of the Fokker–Planck equation, and regarded only the equilibrium-nonequilibrium character of the steady state and the covariant definition of thermodynamic quantities. The bulk of the covariant stochastic thermodynamics, in the spirit of [54], has yet to be developed.

Acknowledgments

The author is thankful to A Bazzani for several discussions over the PhD years, and to M Smerlak for useful comments. The research was partly supported by the National Research Fund Luxembourg in the frame of project FNR/A11/02.

doi:10.1088/1742-5468/2013/07/P07005
Appendix A. Exponential decay to equipartition

We show that average energy has an exponential decay to its steady value, given by the equipartition theorem. We consider energy as defined in equation (22)—setting $m\lambda^2 = 1$—and take its stochastic differential. The key technicality is to use Itô’s lemma

$$dH = \frac{\partial H}{\partial x^a} dx^a + \frac{\partial H}{\partial v^a} dv^a + \frac{1}{2} \frac{\partial^2 H}{\partial v^a \partial v^b} dv^a dv^b,$$

(A.1)

where we only retain contributions of order $dt$. Plugging into the generally covariant Langevin equation we obtain

$$dH = \frac{\partial H}{\partial x^a} v^a dt + \frac{\partial^2 H}{\partial v^a \partial v^b} \delta^{ij} \sigma e^a_i e^b_j dt + \frac{\partial H}{\partial v^a} \left( -\gamma v^a dt - \Gamma_{bc}^a v^b v^c dt + \sqrt{2\sigma} e^a_i dW^i \right).$$

(A.2)

Standard manipulations lead to

$$\dot{H} = -2\gamma \left( H - \frac{n\sigma}{2\gamma} \right) + \sqrt{2\sigma} g_{ab} v^a e^b_i \xi^i.$$

(A.3)

Interestingly, the drift term only depends on the energy itself, so that we can average out the noisy term and obtain a closed equation for the average energy $E_t = \langle H_t \rangle$, with solution

$$E_t = e^{-2\gamma t} \left( E_0 - \frac{n\sigma}{2\gamma} \right) + \frac{n\sigma}{2\gamma}.$$

(A.4)

It decays exponentially to its steady value $n\sigma/2\gamma$, which by the equipartition theorem is identified with $nk_B T/2$, yielding the Einstein relation.

Appendix B. Overdamping limit of metric/diffusion theory

We show that the generally covariant stochastic differential equation (28) (Brownian motion on a manifold) and the generally covariant Langevin equation (19) are equivalent in the overdamping limit $\sigma \to \infty$, at fixed temperature. More precisely, we work with the corresponding diffusion equations, from equations (20)–(27). In fact, we prove the result for the more general theory (G) with independent metric and diffusion matrix discussed in section 5, and specialize the result at the very end. The treatment follows Schuss’s [55, section 8.2]. To simplify, solutions to the perturbative chain of equations that will emerge are given, as they are known from previous literature. A logically close step-by-step derivation can be operated by means of expansions in terms of Hermite polynomials, as is done in [25].

Set both $\beta = 1$ and $(2\pi\hbar)^2/mc^2 = 1$. Define the adimensional energy

$$\phi(x, v) = h_{ab}(x) v^a v^b / 2.$$  

(B.1)

Let $\epsilon = \sigma^{-1}$. We rescale time according to $t = \epsilon^{-1}s$ in the covariant Kramers equation equation (20), and expand to order $\epsilon$ both the generator and the probability density:

$$\left( \epsilon^2 L_2 + \epsilon L_1 + L_0 \right) \left( P_0 + \epsilon P_1 + \epsilon^2 P_2 + \cdots \right) = 0,$$

(B.2)
where

\[ L_0 P = g^{ab} \frac{\partial}{\partial v^a} \left( h_{bc} v^c P + \frac{\partial P}{\partial v^b} \right), \]  
\[ L_1 P = -v^a \frac{\partial P}{\partial x^a} + \frac{\partial}{\partial v^a}(\Omega_{bc}^a v^b v^c P), \]  
\[ L_2 P = -\frac{\partial P}{\partial s}. \]  

We obtain the following expansion:

\[ 0 = L_0 P_0, \]  
\[ 0 = L_0 P_1 + L_1 P_0, \]  
\[ 0 = L_0 P_2 + L_1 P_1 + L_2 P_0, \]  

A solution of the first equation is

\[ P_0(x, v, s) = q_0(x, s) \exp(-\phi(x, v)), \]  

where \( q_0(x, s) \) is an undetermined function. The spatial density is found by performing the integration over velocities:

\[ p_0(x, s) = \int dv P_0(x, v, s) = \frac{Nq_0(x, s)}{\sqrt{h(x)}}. \]  

We also define the conditional probability

\[ \pi(v|x) = \frac{P_0(x, v, s)}{p_0(x, s)} = \frac{\sqrt{h(x)}}{N} \exp(-\phi(x, v)), \]  

which is a multivariate normal distribution. Plugging \( P_0 \) into equation (B.4b), we obtain

\[ g^{ab} \frac{\partial}{\partial v^a} \left( h_{bc} v^c P_1 + \frac{\partial P_1}{\partial v^b} \right) = \left( v^a \frac{\partial}{\partial x^a} \log \frac{q_0}{h} \right) P_0. \]  

A solution is given by

\[ P_1 = (q_1 - v^a \theta_a) P_0, \]  

where \( q_1 = q_1(x, s) \) is again an undetermined function and \( \theta_a \) is easily seen to be

\[ \theta_a = g_{ab} h^{bc} \frac{\partial}{\partial x^c} \log \frac{q_0}{h}. \]  

Substituting into equation (B.4c) and integrating with respect to velocities gives

\[ \frac{\partial p_0}{\partial s} = \frac{\partial}{\partial s} \int P_0 dv = - \int v^a \frac{\partial P_1}{\partial x^a} dv, \]  

where it is assumed that all terms decay sufficiently fast for high velocities, so that

\[ \int dv g^{ab} \frac{\partial}{\partial v^a} \left( h_{bc} v^c + \frac{\partial P_1}{\partial v^b} \right) P_2 = 0, \quad \int dv \frac{\partial}{\partial v^a}(\Omega_{bc}^a v^b v^c P_1) = 0. \]
Equation (B.11) then yields
\[
\frac{\partial p_0}{\partial s} = - \int v^a \frac{\partial}{\partial x^a} \left[ \left( q_1 - v^b \frac{\partial}{\partial x^b} \log \frac{q_0}{q} \right) P_0 \right] dv. \tag{B.13}
\]
The first integrand is odd and its contribution vanishes. In view of equation (B.6), we obtain
\[
\frac{\partial p_0}{\partial s} = \frac{\partial}{\partial x^a} \left[ \left( \int v^a v^b P_0 dv \right) g_{ab} h^{de} \frac{\partial}{\partial x^c} \log \frac{p_0}{\sqrt{h}} \right]. \tag{B.14}
\]
We now perform the integration
\[
\int v^a v^b P_0(x, v, s) dv = p_0(x, s) \int v^a v^b \pi(v|x) dv = p_0(x, s) h^{ab}(x), \tag{B.15}
\]
where we recognize the covariance matrix of \( \pi(\cdot, x) \). Finally, we obtain
\[
\frac{\partial p_0}{\partial s} = \frac{\partial}{\partial x^a} \left( p_0 h^{ab} g_{ad} h^{de} \frac{\partial}{\partial x^c} \log \frac{p_0}{\sqrt{h}} \right). \tag{B.16}
\]
In particular, when \( h_{ab} = g_{ab} \) we obtain equation (27). It follows that in the overdamping limit the generally covariant Langevin equation and Brownian motion on a manifold furnish equivalent descriptions, up to first order in \( \epsilon \).

**Appendix C. Overdamping limit of diffusion with state-dependent temperature and viscosity**

We perform the overdamping limit of theory (F), with state-dependent temperature and viscosity. The particular case with state-dependent temperature is obtained by setting \( h_{ab} = g^{ab} \). The main difference from the above overdamping limit is that the term with Christoffel coefficients disappears, resulting in a minor complication. All goes as before, but for the definition of the operator
\[
L_1 P = -v^a \frac{\partial P}{\partial x^a}. \tag{C.1}
\]
The solution to the first equation in the perturbative chain is the same. As above, plugging \( P_0 \) into equation (B.4b), we obtain
\[
g^{ab} \frac{\partial}{\partial v^a} \left( h_{bc} v^c P_1 + \frac{\partial P_1}{\partial v^b} \right) = \left( v^a \frac{\partial}{\partial x^a} \log q_0 - v^a v^b v^c \Omega_{abc} \right) P_0. \tag{C.2}
\]
We look for solutions in the form
\[
P_1 = \left( q_1 + \frac{1}{3} v^a v^b v^c \Sigma_{abc} - v^a \theta_a \right) P_0, \tag{C.3}
\]
where \( q_1 = q_1(x, s) \) is again a scalar function, \( \theta_a \) is a vector field and \( \Sigma_{abc} \) a tensor. After some work we obtain that \( \Sigma_{abc} = g_{ad} h^{de} \Omega_{ebc} \) and
\[
\theta_a = g_{ab} h^{bc} \left( \frac{\partial}{\partial x^c} \log q_0 - \frac{4}{3} h^{de} \Omega_{dea} - \frac{2}{3} h^{de} \Omega_{ade} \right) = g_{ab} h^{bc} \left( \frac{\partial}{\partial x^c} \log q_0 - \frac{2}{3} \partial_c \log \sqrt{h} + \frac{2}{3} h_{ad} \partial_d h^{de} \right). \tag{C.4}
\]
Generally covariant state-dependent diffusion satisfy equation (C.2), where we have used the identities

\[ h^{bc} \Omega_{bca} = \partial_a \log \sqrt{h}, \quad h^{bc} \Omega_{abc} = -h_{ac} \partial_b h^{bc} - \partial_a \log \sqrt{h}. \]  

(C.5)

Substituting into equation (B.4c) and integrating with respect to \( dv \) we obtain

\[ \frac{\partial p_0}{\partial s} = -\frac{\partial}{\partial x^a} \int dv \left(-v^a v^b \theta_b + \frac{1}{3} \partial^c h_{abc} h^{[e} \Omega_{ecd]} \right) P_0 = -\frac{\partial}{\partial x^a} (j_1^a + j_2^a), \]  

(C.6)

where again the term with \( q_1 \) drops for it gives an odd integrand. Given equation (B.15), the first current contribution reads

\[ j_1^a = p_0 h^{ab} g_{bc} \left(-\partial^c \log p_0 - \frac{1}{3} \partial^c \log \sqrt{h} - \frac{2}{3} \partial_d h^{dc} \right), \]  

(C.7)

where we have raised indices with \( h^{ab} \). By the Isserli–Wick theorem

\[ \int dv v^a v^b v^c v^d P_0 = p_0 \left(h^{ab} h^{cd} + h^{ac} h^{bd} + h^{ad} h^{bc}\right). \]  

(C.8)

Letting \( G^{ab} = h^{ac} g_{cd} h^{db} \)

\[ j_2^a = \frac{p_0}{3} \left[h^{ab} g_{bc} (-\partial_d h^{dc} - \partial^e \log \sqrt{h}) + 2 h^{ab} G^{cd} \Omega_{cd} \right], \]  

(C.9)

for a total Fokker–Planck current

\[ j^a = -h^{ab} g_{bc} \left[\partial^c p_0 + \left(\frac{2}{3} \partial^c \log \sqrt{h} + \partial_d h^{dc}\right) p_0 \right] + \frac{2}{3} h^{ab} G^{cd} \Omega_{cd} p_0. \]  

(C.10)

In particular when \( h^{ab} = g^{ab} \) we obtain the Itô generator \( \partial_s p_0 = -\partial_d \partial_a (g^{ab} p_0) \), while in the one-dimensional case we find that

\[ \frac{\partial p_0}{\partial s} = -\frac{g}{h^2} \left[-\partial p_0 + \frac{p_0}{3} \left(2 + \frac{g}{h}\right) \partial \log h \right]. \]  

(C.11)

Appendix D. Comparison of models

The following diagram summarizes the interdependence of the several Langevin equations that we have considered. The symbol \( \rightsquigarrow \) means ‘is replaced by’.

\[ \text{(D.1)} \]
Generally covariant state-dependent diffusion

Table D.1. State-dependent diffusion theories in second-order formalism.

| Model                      | References | Langevin equation                                               | Einstein relation | Steady state |
|----------------------------|------------|-----------------------------------------------------------------|-------------------|--------------|
| A  Generally covariant     | here, [38, 39] | \( \ddot{x}^a + \Gamma^a_{bc} \dot{x}^b \dot{x}^c = -\gamma \dot{x}^a + \sqrt{2}\sigma e^a_i \zeta^i \) | \( \gamma = \beta \sigma \) | \( g \exp[-(\beta/2)g_{ab}v^av^b] \) |
| B  Covariant (Graham)      | [10, 11]   | \( \ddot{x}^a = e^a_b \partial_t x^b \dot{x}^c - \gamma \dot{x}^a + \sqrt{2}\sigma e^a_i \zeta^i \) | \( \gamma = \beta \sigma \) | ?            |
| C  State dep. temperature  | [24]--[28] | \( \ddot{x}^a = -\gamma \dot{x}^a + \sqrt{2}\sigma e^a_i \zeta^i \) | \( \gamma = \beta \sigma \) | ?            |
| D  State dep. viscosity    | [19, 20, 46, 47] | \( \ddot{x}^a = -\gamma_b^a \dot{x}^b + \sqrt{2}\sigma e^a_i \zeta^i \) | \( \gamma_b^a = \beta \sigma g^{ac} \delta_{cb} \) | \( \exp[-(\beta/2)\delta_{ab}v^av^b] \) |
| E  Pure posticipatory SDE  | [5, section V] | —                                                               | —                 | —            |
| F  St. dep. visc. and temp.| [49, 50]   | \( \ddot{x}^a = -\gamma_b^a \dot{x}^b_t + \sqrt{2}\sigma e^a_i \zeta^i \) | \( \gamma_b^a = \beta \sigma g^{ac} h_{cb} \) | ?            |
| G  Metric/diffusion        | here       | \( \ddot{x}^a + \Omega^a_{bc} \dot{x}^b \dot{x}^c = -\gamma_b^a \dot{x}^b_t + \sqrt{2}\sigma e^a_i \zeta^{ia} \) | \( \gamma_b^a = \beta \sigma g^{ac} h_{cb} \) | \( h \exp[-(\beta/2)h_{ab}v^av^b] \) |

\( \Omega^a_{bc} \) are the Christoffel coefficients of \( h^{ab} \).
| Smoluchovski eq. | FP current | 1D SS<sup>a</sup> | Gauge inv. | Covariance | DB<sup>b</sup> |
|------------------|------------|-------------------|------------|------------|------------|
| A                | $\dot{x}^a = -g^{bc} \Gamma^a_{bc} + \sqrt{2} e_i^a \xi^i$ | $-g^{ab} (\partial_b p - p \partial_b \log \sqrt{g})$ | $\sqrt{g}$ | Yes | Yes | Yes |
| B                | $\dot{x}^a = \delta^{ij} e_i^b \partial_b e_j^a + \sqrt{2} e_i^a \xi^i$ | $-g^{ab} \partial_b p - (\delta^{ij} e_i^a \partial_b e_j^b) p$ | $\sqrt{g}$ | No | Yes | No |
| C                | $\dot{x}^a = \sqrt{2} e_i^a \xi^i$ | $-\partial_b (g^{ab} p)$ | $g$ | Yes | No | No |
| D                | $\dot{x}^a = \partial_b g^{ab} + \sqrt{2} e_i^a \xi^i$ | $-g^{ab} \partial_b p$ | Const. | Yes | No | Yes |
| E                | (Omitted) | $\dot{x}^a = 2 \delta^{ij} e_i^b \partial_b e_j^a + \sqrt{2} e_i^a \xi^i$ | $-h^{ab} g_{bc} [\partial^c p + \left(2 \partial^c \log \sqrt{h} + \partial_d h^{de} \right) p]$ | $h^{2/3} e^f g^{ab} \Omega_{cde}^b$ | Yes | No | No |
| F                | $\dot{x}^a = -g^{bc} \Gamma^a_{bc} - \partial_b \log \sqrt{h/g} + \sqrt{2} e_i^a \xi^i$ | $-g^{ab} \left(\partial_b p - p \partial_b \log \sqrt{h}\right)$ | $\sqrt{h}$ | Yes | Yes | Yes |

<sup>a</sup> Steady state in the one-dimensional case.

<sup>b</sup> Detailed balance, independently of the specific choice of metric, coordinates and gauge. Specific models might obey detailed balance.

<sup>c</sup> $G^{ab} = h^{ac} g_{cd} h^{db}$. 

---

Generally covariant state-dependent diffusion
References

[1] Sekimoto K, *Langevin equation and thermodynamics*, 1998 Prog. Theor. Phys. 130 (Suppl.) 17
[2] Seifert U, *Stochastic thermodynamics: principles and perspectives*, 2008 Eur. Phys. J. B 64 423
[3] Van den Broeck C and Esposito M, *Three faces of the second law. II. Fokker–Planck formulation*, 2010 Phys. Rev. E 82 011144
[4] Hänggi P, Talkner P and Borkovec M, *Reaction-rate theory: 50 years after Kramers*, 1990 Rev. Mod. Phys. 62 251
[5] Lau A W C and Lubensky T C, *State-dependent diffusion: thermodynamic consistency and its path integral formulation*, 2007 Phys. Rev. 76 011123
[6] Klimontovich Y L, *Nonlinear Brownian motion*, 1994 Phys.—Usp. 37 737
[7] Hsu E P, 2002 Stochastic Analysis on Manifolds (Providence, RI: American Mathematical Society)
[8] Stroock D W, 2000 *An Introduction to the Analysis of Paths on a Riemannian Manifold* (Providence, RI: American Mathematical Society)
[9] Van Kampen N G, *Brownian motion on a manifold*, 1986 J. Stat. Phys. 44 1
[10] Graham R, *Covariant formulation of non-equilibrium statistical thermodynamics*, 1977 Z. Phys. B 26 397
[11] Muratore-Ginanneschi P, *Path integration over closed loops and Gutzwiller’s trace formula*, 2003 Phys. Rep. 383 299
[12] Obata T, Hara H and Endo K, *Differential geometry of nonequilibrium processes*, 1992 Phys. Rev. A 45 6097
[13] Crooks G E, *Measuring thermodynamic length*, 2007 Phys. Rev. Lett. 99 100602
[14] Ren J, Hänggi P and Li B, *Berry-phase-induced heat pumping and its impact on the fluctuation theorem*, 2010 Phys. Rev. Lett. 104 170601
[15] Chernyak V Y, Klein J R and Sinitsyn N A, *Algebraic topology and the quantization of fluctuating currents*, 2011 arXiv:1204.2011
[16] Jiang D-J, Qian M and Qian M-P, 2004 *Mathematical Theory of Nonequilibrium Steady States* (Berlin: Springer)
[17] Polettini M, *Nonequilibrium thermodynamics as a gauge theory*, 2012 Eur. Phys. Lett. 97 30003
[18] Polettini M, *Diffusion in nonuniform temperature and its geometric analog*, 2013 Phys. Rev. E 87 032126
[19] Sancho J M, San Miguel M and Dür D, *Adiabatic elimination for systems of Brownian particles with nonconstant damping coefficients*, 1982 J. Stat. Phys. 28 291
[20] Sancho J M, *Brownian colloidal particles: Ito, Stratonovich or a different stochastic interpretation*, 2011 Phys. Rev. E 84 062102
[21] Lançon P, Batrouni G, Lobry L and Ostrowsky S, *Drift without flux: Brownian walker with a space dependent diffusion coefficient*, 2001 Europhys. Lett. 54 28
[22] Volpe G, Helden L, Brettschneider T, Wehr J and Bechinger C, *Influence of noise on force measurements*, 2010 Phys. Rev. Lett. 104 170602
[23] Volpe G, Helden L, Brettschneider T, Wehr J and Bechinger C, *Force measurement in the presence of Brownian noise: equilibrium-distribution method versus drift method*, 2011 Phys. Rev. E 83 041113
[24] Mannella R and McClintock P V E, *Comment on influence of noise on force measurements*, 2011 Phys. Rev. Lett. 107 078901
[25] Mannella R et al, *Reply*, 2011 Phys. Rev. Lett. 107 078902
[26] Sekimoto K, 2010 *Stochastic Energetics* (Springer Lecture Notes in Physics vol 799) (Berlin: Springer)
[27] Matsuo M and Sasa S, *Stochastic energetics of nonuniform temperature systems*, 2000 Physica A 276 188
[28] Bringer E and Bourdon A, *Colloid transport in non-uniform temperature*, 2003 Phys. Rev. E 67 011404
[29] Celani A, Bo S, Eichhorn R and Aurell E, *Anomalous thermodynamics at the micro-scale*, 2012 arXiv:1206.1742
[30] Van Kampen N G, *Relative stability in nonuniform temperature*, 1988 IBM J. Res. Dev. 32 107
[31] Landauer R, *Motion out of noisy states*, 1988 J. Stat. Phys. 52 233
[32] Van Kampen N G, *Ito vs. Stratonovich*, 1981 J. Stat. Phys. 24 175
[33] Klimontovich Yu L, *Ito, Stratonovich and kinetic forms of stochastic equations*, 1990 Physica A 163 515
[34] Sokolov I M, *Ito, Stratonovich, Hänggi and all the rest: the thermodynamics of interpretation*, 2010 Chem. Phys. 375 359
[35] Yuan R and Ao P, *Beyond Ito vs. Stratonovich*, 2012 J. Stat. Mech. P07010
[36] Mannella R and McClintock P V E, *Ito vs. Stratonovich: thirty years later*, 2012 Fluct. Noise Lett. 11 1240010
[37] Cartan É, *Sur les variétés à connexion affine, et la théorie de la relativité généralisée (première partie)*, 1923 Ann. Sci. Ec. Norm. Sup. 40 325

doi:10.1088/1742-5468/2013/07/P07005
Generally covariant state-dependent diffusion

[36] Carroll S, 2003 Spacetime and Geometry: An Introduction to General Relativity (Boston, MA: Addison-Wesley)

[37] Muratore-Ginanneschi P, On the use of stochastic differential geometry for non-equilibrium thermodynamic modeling and control, 2013 J. Phys. A: Math. Theor. 46 275002

[38] Kleinert H and Shabanov S V, Theory of Brownian motion of a massive particle in spaces with curvature and torsion, 1998 J. Phys. A: Math. Gen. 31 7005

[39] Castañeda-Priego R, Castro-Villarreal P, Estrada-Jiménez S and Méndez-Alcaraz J M, Brownian motion of free particles on curved surfaces, 2012 arXiv:1211.5799

[40] Batrouni G G, Kawai H and Rossi P, Coordinate-independent formulation of the Langevin equation, 1986 J. Math. Phys. 27 1646

[41] Castro Villarreal P, Brownian motion meets Riemann curvature, 2010 J. Stat. Mech. P08006

[42] Tubber P F and Yang Y, A paradox of state dependent diffusion and how to resolve it, 2012 arXiv:1204.1590

[43] Berger J, The case for objective bayesian analysis, 2006 Bayesian Anal. 1 385

[44] Van den Broeck C and Esposito M, Three faces of the second law. II. Fokker–Planck formulation, 2010 Phys. Rev. E 82 011144

[45] Rumpf H, Stochastic quantization of Einstein gravity, 1986 Phys. Rev. D 33 942

[46] Hasegawa H, Classical open systems with nonlinear nonlocal dissipation and state-dependent diffusion: dynamical responses and the Jarzynski equality, 2011 Phys. Rev. E 84 051124

[47] Arnold P, Langevin equations with multiplicative noise: resolution of time discretization ambiguities for equilibrium systems, 2000 Phys. Rev. E 61 6091

[48] Arnold P, Symmetric path integrals for stochastic equations with multiplicative noise, 2000 Phys. Rev. E 61 6099

[49] Van Kampen N G, Diffusion in inhomogeneous media, 1987 Z. Phys. B 68 135

[50] Jayannavar A M and Mahato M C, Macroscopic equation of motion in inhomogeneous media: a microscopic treatment, 1995 Pramana J. Phys. 45 369

[51] Chacón-Acosta G and Kremer G, Fokker–Planck-type equations for a simple gas and for a semirelativistic Brownian motion from a relativistic kinetic theory, 2007 Phys. Rev. E 76 021201

[52] Smerlak M, Tailoring diffusion in analogue spacetimes, 2012 Phys. Rev. E 85 041134

[53] Girolami M and Calderhead B, Riemann manifold Langevin and Hamiltonian Monte Carlo methods, 2011 J. R. Stat. Soc. B 73 123

[54] Imparato A and Peliti L, Fluctuation relations for a driven Brownian particle, 2006 Phys. Rev. E 74 026106

[55] Schuss Z, 2010 Theory and Applications of Stochastic Processes (New York: Springer)

doi:10.1088/1742-5468/2013/07/P07005