# Dynamics and kinematics of reciprocal diffusions

Bernard C. Levy
*Department of Electrical and Computer Engineering, University of California, Davis,
California 95616*

Arthur J. Krener
*Department of Mathematics, University of California, Davis, California 95616*

This paper examines the dynamics and kinematics of reciprocal diffusions. Reciprocal processes were introduced by Bernstein in 1932, and were later studied in detail by Jamison. The reciprocal diffusions are constructed here by specifying their finite joint densities in terms of the Green's function of a general heat operator, and an end-point density. A path integral interpretation of the heat operator Green's function is provided, which is used to derive a stochastic form of Newton's law, as well as a conditional distribution for the velocity of a diffusing particle given its position. These results are then employed to derive two conservation laws expressing the conservation of mass and momentum. The conservation laws do not form a closed system of equations, in general, except for two subclasses of reciprocal diffusions, the Markov and quantum diffusions.

## I. INTRODUCTION

Reciprocal processes find their origin in an attempt of Schrödinger[1] to develop a stochastic formulation of quantum mechanics in terms of Markov diffusions. In this formulation, the probability density of a particle at a given time is represented as the product of two functions satisfying respectively forward and backward Fokker–Planck equations. However, two aspects of Markov diffusions created serious obstacles to Schrödinger's investigation. The first is that, unlike the equations of classical or quantum mechanics, Markov diffusion models are not time symmetric, since they usually involve different forward and backward drifts. Also, the stochastic differential equations describing Markov diffusions are *first-order* equations. As such they model accurately the kinematics of a diffusing particle subjected to random fluctuations, but they do not capture dynamical effects.

The lack of time symmetry of Markov diffusions motivated Bernstein[2] in 1932 to introduce reciprocal processes. The reciprocity property is essentially one of time locality, where given a fixed interval, the process over the interval depends only on its values at both ends of the interval, and on the excitations to which it is subjected in the intervening times. The class of reciprocal processes contains Markov processes, as well as other processes. In fact, it is interesting to observe that the Markov random fields in the sense of P. Lévy[3] reduce in one dimension to reciprocal, not Markov, processes, so that they should probably have been called reciprocal fields. Reciprocal processes were subsequently studied in detail by Jamison,[4] who showed that they could be constructed from Markov processes by first pinning a Markov process at both ends of a fixed time interval, and then assigning an arbitrary probability distribution to the end points of the process.

In parallel with these developments, a systematic study of the dynamical properties of Markov diffusions was undertaken by Nelson,[5] which ultimately evolved into a complete stochastic interpretation of quantum mechanics, called stochastic mechanics.[6] The starting point of Nelson's investigation was the observation that to model the dynamics of a diffusing praticle, one must be able to define the concept of acceleration in a stochastic context. This was accomplished by working simultaneously with the forward and backward models of a Markov

diffusion, and defining the acceleration as an average of mean forward and backward derivatives.

However, this definition of the stochastic acceleration is rather unintuitive. This in part motivated Krener[7,8] to propose second-order models of reciprocal diffusions, where a reciprocal diffusion is defined as one which is constructed by applying Jamison's pinning procedure to a Markov diffusion. Krener showed that reciprocal diffusions satisfy locally a stochastic version of Newton's law, which was then employed to derive conservation laws for the probability density and momentum density of a reciprocal diffusion. The derivation of Newton's law in Ref. 8 relied on a characterization of Jamison[4] for the three-point transition density of a reciprocal process in terms of the two-point transition density of a Markov process, and on a short-time asymptotic expansion of the transition density of Markov diffusions. Subsequently, it was shown by Clark[9] that for reciprocal diffusions in $\mathbf{R}^n$, the forces appearing in Newton's law can be described in terms of $n + n(n-1)/2$ quantities, called the "reciprocal invariants" of the process. In addition, Clark extended to reciprocal diffusions the change of measure formula developed by Girsanov for Markov diffusions.[10]

Note that although reciprocal diffusions, which are often called Bernstein diffusions, have been used extensively in recent years by Zambrini[11] and other researchers[12] in the context of stochastic mechanics, the class of diffusions that they consider corresponds actually to the Markov diffusions of Schrödinger, which is only a subclass of the full set of reciprocal diffusions discussed here.

The goal of the present paper is to demonstrate that to any dynamical system, described by its Hamiltonian, we can associate a class of reciprocal diffusions whose dynamics and kinematics can then be interpreted in terms of the underlying classical system. The procedure used to go from the classical system to the corresponding reciprocal class can be viewed as a "stochastic quantization" method, although this term has been used with a different meaning in other contexts. The procedure we employ consists in associating to each Hamiltonian parametrized by a scalar and covector potential pair a generalized heat operator, corresponding to a Markov diffusion with creation and killing. Then, Jamison's construction is employed to characterize the finite joint densities of the corresponding reciprocal class in terms of the heat operator Green's function and a positive measure for the end points of the diffusion. In this context, it is shown that the joint densities of the diffusion are not affected by the application of a gauge transformation to the scalar and covector potentials, and the end-point measure. Also, when the gauge is fixed, the scalar and covector potentials define a set of reciprocal invariants for the diffusion.

To obtain a dynamical characterization of reciprocal diffusions, we derive a path integral representation of the heat operator Green's function used to construct their finite joint densities. Unlike the usual Feynman path integrals, the summation does not involve complex probability "amplitudes," but real probabilities. It is therefore similar to the probabilistic interpretation of Feynman integrals given by Kac.[13] The path integral representation yields a short-time asymptotic of the heat operator Green's function, which is then used to derive the stochastic Newton law introduced by Krener, and to show that the distribution of the empirical velocity of a diffusing particle given its position is approximately Gaussian. These results are used to derive two conservation laws for the probability density and momentum density of reciprocal diffusions, which take a form identical to the equations of fluid mechanics. These two conservation laws are in fact part of an infinite chain of conservation laws which do not form a closed set of equations, in general, except for two subclasses of reciprocal diffusions. The first subclass is of course the Markov diffusions, in which case the two conservation laws reduce to forward and backward propagating Fokker–Planck equations. The second class, which seems to have been overlooked until now, consists of what we shall call the *quantum diffusions*, for which the two conservation laws are equivalent to Schrödinger's equation. These diffusions form the basis for an alternative version of stochastic mechanics, described in Ref. 14.

The paper is organized as follows. In Sec. II, the joint probability densities of a reciprocal diffusion are characterized in terms of the Green's function of a generalized heat operator and an end point measure. The Markov case corresponds to the situation where the end point measure is factorable. A path integral representation of the heat operator Green's function is obtained in Sec. III. This representation is employed in Sec. IV to derive the stochastic Newton law of reciprocal diffusions. A simplified reciprocal Girsanov transformation is presented in Sec. V. In Sec. VI, it is shown that the distribution for the empirical velocity of a diffusing particle given its position is approximately Gaussian, and is parametrized by a mean velocity vector and a stress tensor. The conservation laws of reciprocal diffusions are obtained in Sec. VII. Unfortunately, these conservation laws do not close, in general, except for the subclasses of Markov and quantum diffusions, where the mean velocity and stress tensor satisfy algebraic constraints that provide two possible closure rules for the reciprocal conservation laws, which are discussed in Sec. VIII.

## II. RECIPROCAL AND MARKOV DIFFUSIONS

As starting point, we recall that a process $x(t) \in \mathbf{R}^n$ defined for $t \in [0,T]$ is reciprocal if for arbitrary subintervals $[s,t]$ of $[0,T]$, the values of $x(.)$ in the interior and exterior of $[s,t]$ are conditionally independent given the end values $x(s)$ and $x(t)$. From this definition it is clear that if a process is Markov, it is necessarily reciprocal, but the converse is not true.

The space $\mathbf{R}^n$ is endowed with a time-dependent, but space-independent, metric $\sigma_{ij}(t)$ which is assumed uniformly positive definite over $[0,T]$. This means that if the matrix $\sigma^{ij}(t)$ denotes the inverse of the metric, i.e.,

$$\sigma^{ik}(t)\sigma_{kj}(t) = \delta^i_j = \begin{cases} 1 & \text{if} \quad i=j \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

there exists positive constants $a_0$ and $a_1$ such that if $v$ denotes an arbitrary vector of $\mathbf{R}^n$,

$$a_0|v|^2 < \sigma^{ij}(t)v_i v_j < a_1|v|^2 \tag{2.2}$$

for all $t \in [0,T]$. We consider a dynamical system with Hamiltonian

$$H(x,p,t) = \tfrac{1}{2}\sigma^{ij}(t)(p_i - A_i(x,t))(p_j - A_j(x,t)) + \phi(x,t), \tag{2.3}$$

where we have used the standard tensor contraction convention, with repeated upper and lower indices corresponding to a summation. In Eq. (2.3), $p_i$ is the generalized momentum of the system, and the scalar and covector potentials $\phi(x,t)$ and $A_i(x,t)$ are given.

To the Hamiltonian $H(x,p,t)$ we can associate an elliptic operator $\mathbf{H}$ obtained by replacing $p_j$ by $-\nabla_j$ where $\nabla_j$ denotes the differentiation with respect to $x^j$. This substitution can be viewed as the stochastic analog of the quantization rule $p_j \leftrightarrow -i\nabla_j$ of quantum mechanics. The resulting operator $\mathbf{H}$ is given by

$$\mathbf{H} = \tfrac{1}{2}(\nabla^i + A^i)(\nabla_i + A_i) + \phi \tag{2.4a}$$

$$= \tfrac{1}{2}\Delta + A^i\nabla_i + \tfrac{1}{2}(\nabla^i A_i + A^i A_i) + \phi, \tag{2.4b}$$

where $\Delta$ is the Laplace operator

$$\Delta \triangleq \sigma^{ij}(t)\frac{\partial^2}{\partial x^i \partial x^j} \tag{2.5}$$

expressed with respect to the metric $\sigma^{ij}$. Now, consider the parabolic operator

$$\mathbf{L} = \mathbf{H} - \frac{\partial}{\partial t} \qquad (2.6)$$

and the Green's function $G(x,s;y,t)$ given by

$$\mathbf{L}_{y,t} G(x,s;y,t) = 0, \quad t \geqslant s, \qquad (2.7a)$$

$$G(x,s;y,s) = \delta(y - x), \qquad (2.7b)$$

where the subscripts $\{y,t\}$ indicate $\mathbf{L}$ is acting on the variables $y$ and $t$. It is also required that the Green's function $G(x,s;y,t)$ should decay as $y \to \infty$. When $\sigma_{ij}$, $A_i$, $\partial A_i / \partial x^k$ and $\phi$ are bounded and satisfy a uniform Hölder continuity condition over $\mathbf{R}^n \times [0,T]$, it is shown in Ref. 15 that $G$ exists and is unique. Furthermore, $G(x,s;y,t)$ is nonnegative and satisfies

$$G(x,s;y,t) \leqslant C_1 (t-s)^{-n/2} \exp - \frac{a_0' |y-x|^2}{4(t-s)}, \qquad (2.8a)$$

$$\left| \frac{\partial G(x,s;y,t)}{\partial y^i} \right| \leqslant C_2 (t-s)^{-(n+1)/2} \exp - \frac{a_0' |y-x|^2}{4(t-s)} \qquad (2.8b)$$

for any $a_0' < a_0$, where $C_1$ and $C_2$ are constants.

The Green's function $G(x,s;y,t)$ has a number of useful properties. The most notable is the transition property

$$G(x,s;z,u) = \int G(x,s;y,t) G(y,t;z,u) dy \qquad (2.9)$$

for $s < t < u$. Like the Green's function of standard Markov diffusion operators, $G(x,s;y,t)$ can be interpreted as a transition probability density, since it is positive and satisfies the initial condition (2.7b). However, this density is *unnormalized*. To see this, let

$$Z_R(x,s;t) = \int_{B(x,R)} G(x,s;y,t) dy, \qquad (2.10)$$

where $B(x,R)$ is the ball of radius $R$ centered at $x$. Integrating Eq. (2.7a) over $B(x,R)$ and using the divergence theorem, we find

$$\frac{\partial}{\partial t} Z_R(x,s;t) = \int_{S(x,R)} \left( \frac{\nabla^i G}{2} + A^i G \right) n_i \, dS + \int_{B(x,R)} ((A^i A_i - \nabla^i A_i)/2 + \phi) G \, dy, \quad (2.11)$$

where $S(x,R)$ is the sphere of radius $R$ centered at $x$, $dS$ is an elementary surface element, and $n$ is a unit vector perpendicular to the surface. Letting $R \to \infty$ inside (2.11), and using the bounds (2.8a)–(2.8b) as well as the boundedness of $A_i$, we see that the surface integral vanishes, so that the integral $Z(x,s;t)$ of $G(x,s;y,t)$ over $\mathbf{R}^n$, i.e., the total probability mass at time $t$ of the density $G(x,s;y,t)$, evolves according to

$$\frac{\partial}{\partial t} Z(x,s;t) = \int G(x,s;y,t) c(y,t) dy, \qquad (2.12)$$

where

$$c(x,t) \triangleq \tfrac{1}{2}(A^i A_i - \nabla^i A_i)(x,t) + \phi(x,t) \qquad (2.13)$$

denotes the rate of mass creation at point $x$ and time $t$. Note that depending on whether $c(x,t) > 0$ or $c(x,t) < 0$, there is actually mass creation or annihilation. This shows that $Z(x,s;t) \neq 1$, so that unless $c(x,t) \equiv 0$, the probability mass is not conserved.

In the special case when $c(x,t) = 0$,

$$\mathbf{L} = \frac{1}{2} \frac{\partial^2}{\partial x^i \partial x^j} (\sigma^{ij} \cdot) + \frac{\partial}{\partial x^i} (A^i \cdot) - \frac{\partial}{\partial t} \qquad (2.14)$$

is the forward operator for a standard Markov diffusion with diffusion matrix $\sigma^{ij}(t)$ and drift $b^i(x,t) = -A^i(x,t)$.

Another property of the Green's function $G$ involves its backward propagation in time. Consider the adjoint operators

$$\mathbf{H}^* = \tfrac{1}{2}\Delta - A^i \nabla_i + \tfrac{1}{2}(-\nabla^i A_i + A^i A_i) + \phi, \qquad (2.15a)$$

$$\mathbf{L}^* = \mathbf{H}^* + \frac{\partial}{\partial t}, \qquad (2.15b)$$

where we note that $\mathbf{H}^*$ is obtained by replacing $A^i$ by $-A^i$ in $\mathbf{H}$. Then, for $s \leqslant t$, the Green's function $G(x,s;y,t)$ satisfies the backward equation

$$\mathbf{L}^*_{x,s} G(x,s;y,t) = 0, \qquad (2.16a)$$

$$G(x,t;y,t) = \delta(x - y). \qquad (2.16b)$$

We define reciprocal and Markov diffusions by specifying their finite joint densities as follows.

*Definition 2.1*: A process $x(t)$ is a reciprocal diffusion over $[0,T]$ if given arbitrary times $t_0 = 0 < t_1 < \cdots < t_N = T$, the joint probability density of $x(t_0)$, $x(t_1)$, ... , $x(t_N)$ is given by

$$p(x_0,0;x_1,t_1;\ldots;x_N,T) = q(x_0,x_N) \prod_{k=0}^{N-1} G(x_k,t_k;x_{k+1},t_{k+1}), \qquad (2.17)$$

where $G$ is the Green's function of a heat operator of the form (2.4)–(2.5) and $q(x_0,x_N)$ is a positive density such that

$$\int \int q(x_0,x_N) G(x_0,0;x_N,T) dx_0 \, dx_N = 1. \qquad (2.18)$$

When $q(x_0,x_N)$ can be factored as

$$q(x_0,x_N) = q_f(x_0) q_b(x_N) \qquad (2.19)$$

$x(t)$ is said to be a Markov diffusion.

From the above definition, we see that when the interval of definition of a reciprocal diffusion is restricted from $[0,T]$ to a subinterval $[s,t]$, the density $q(x,s;y,t)$ that needs to be applied to the new end points $(s,t)$ is given by

$$q(x,s;y,t) = \int \int q(x_0,x_T) G(x_0,0;x,s) G(y,t;x_T,T) dx_0 \, dx_T . \qquad (2.20)$$

The expression (2.20) indicates that $q(x,s;y,t)$ with $s \leqslant t$ satisfies the two evolution equations

$$\mathbf{L}_{y,q}(x,s;y,t) = 0,$$
(2.21a)

$$\mathbf{L}^*_{x,s}q(x,s;y,t) = 0.$$
(2.21b)

The specification (2.17)–(2.18) of a reciprocal diffusion implies also that for $s < t < u$, the three-point transition density $r(x,s;y,t;z,u)$ of $x(t)$, i.e., the conditional density of $x(t)$ given the outer values $x(s)$ and $x(u)$, can be expressed as

$$r(x,s;y,t;z,u) = \frac{G(x,s;y,t)G(y,t;z,u)}{G(x,s;z,u)},$$
(2.22)

which, except for the fact that $G(x,s;y,t)$ is an unnormalized density, satisfies precisely the relation obtained by Jamison[4] linking the three-point transition density of a reciprocal process to the two-point transition-density of a Markov process.

Finally, observe that several potential pairs $\{A_i(x,t),\phi(x,t)\}$ and end-point densities $q(x_0,x_T)$ may yield the same reciprocal diffusion. To see this, let $\chi(x,t)$ be an arbitary differentiable function, and consider the *gauge transformation*

$$G'(x,s;y,t) = \exp(\chi(x,s))G(x,s;y,t)\exp(-\chi(y,t)),$$
(2.23a)

$$q'(x_0,x_T) = \exp(-\chi(0,x_0))q(x_0,x_T)\exp(\chi(T,x_T)).$$
(2.23b)

It is easy to verify that under this gauge transformation, the joint densities (2.17) and the three-point conditional $r(x,s;y,t;z,u)$ are invariant. Furthermore, $G'(x,s;y,t)$ is the Green's function of a heat operator with potentials

$$A_i'(x,t) = A_i(x,t) - \frac{\partial \chi}{\partial x^i}(x,t),$$
(2.24a)

$$\phi'(x,t) = \phi(x,t) + \frac{\partial \chi}{\partial t}(x,t).$$
(2.24b)

This motivates the following definition and lemma.

*Definition 2.2*: For a fixed metric $\sigma_{ij}$, the triples $\{A_i(x,t),\phi(x,t),q(x_0,x_T)\}$ and $\{A_i'(x,t),\phi'(x,t),q'(x_0,x_T)\}$ are equivalent if they parametrize the same reciprocal diffusion.

*Lemma 2.1*: Two triples $\{A_i,\phi,q\}$ and $\{A_i',\phi',q'\}$ are equivalent if and only if they are related by a gauge transformation of the form (2.23b) and (2.24).

*Proof*: If two triples are equivalent, their Green's functions $G(x,s;y,t)$ and $G'(x,s;y,t)$ yield the same three-point density $r(x,s;y,t;z,u)$ for $s < t < u$, so that

$$\frac{G(x,s;y,t)G(y,t;z,u)}{G(x,s;z,u)} = \frac{G'(x,s;y,t)G'(y,t;z,u)}{G'(x,s;z,u)}.$$
(2.25)

Define

$$R(x,s;y,t) \triangleq \frac{G'(x,s;y,t)}{G(x,s;y,t)},$$
(2.26)

where we observe that in the limit as $t \to s$ and $y \to x$,

$$R(x,s;y,t) \to 1. \tag{2.27}$$

The relation (2.25) implies

$$R(x,s;z,u) = R(x,s;y,t)R(y,t;z,u), \tag{2.28}$$

and differentiating both sides of Eq. (2.28) with respect to $y^i$, we find

$$\frac{\partial}{\partial y^i} \ln R(x,s;y,t) = -\frac{\partial}{\partial x^i} \ln R(y,t;z,u). \tag{2.29}$$

Since the left-hand side of Eq. (2.29) does not depend on $(z,u)$ and the right-hand side on $(x,s)$, it is a function of $(y,t)$ only, from which we can immediately deduce that

$$\ln R(x,t;y,s) = -\chi(y,t) + \lambda(x,s) \tag{2.30}$$

for some functions $\chi(y,t)$ and $\lambda(x,s)$. From the limit (2.27) we find $\chi(x,s) = \lambda(x,s)$, so that $G'(x,s;y,t)$ and $G(x,s;y,t)$ are related by a gauge transformation, which in turn implies that the end-point densities must satisfy Eq. (2.23b). $\qquad\square$

The reciprocal and Markov diffusions introduced in Definition 2.1 are specified in terms of the potentials $\{A_i(x,t),\phi(x,t)\}$ and density $q(x_0,x_T)$. The potentials $\{A_i,\phi\}$ are determined from physical considerations. On the other hand, the end-point density $q(x_0,x_T)$ must be computed from the joint probability density $p(x_0,0;x_T,T)$ of $x(0) = x_0$ and $x(T) = x_T$, which provides the most natural information about the boundary conditions of the diffusion. From Eq. (2.17), we see that $q(x_0,x_T)$ can be obtained from $p(x_0,0;x_T,T)$ by using

$$p(x_0,0;x_T,T) = G(x_0,0;x_T,T)q(x_0,x_T). \tag{2.31}$$

The Markov case is more complicated, since unlike the reciprocal case, where the joint probability density $p(x_0,0;x_T,T)$ can be assigned arbitrarily, the density $p(x_0,0;x_T,T)$ has some internal structure. In this case, it is more reasonable to assume that the marginal densities $\rho(x_0,0)$ and $\rho(x_T,T)$ of $x(0)$ and $x(T)$ are given. According to Eq. (2.31), the end-point densities $q_f(x_0)$ and $q_b(x_T)$ must be related to $\rho(x_0,0)$ and $\rho(x_T,T)$ through the coupled integral equations

$$\rho(x_0,0) = q_f(x_0) \int G(x_0,0;x_T,T)q_b(x_T)dx_T, \tag{2.32a}$$

$$\rho(x_N,T) = q_b(x_T) \int G(x_0,0;x_T,T)q_f(x_0)dx_0, \tag{2.32b}$$

which were first introduced by Schrödinger.[1] Multiplying $q_f(x_0)$ by an arbitrary constant, and dividing $q_b(x_T)$ by the same constant leaves (2.32a) and (2.32b) unchanged, so that the solutions of these equations will only be fixed up to an arbitrary scaling. When such a scaling is provided, the existence and uniqueness of such solutions was studied in Refs. 16 and 2.

From Eqs. (2.17) and (2.20), we see that the marginal density $\rho(x,t)$ of a reciprocal diffusion $x(t)$ satisfies

$$\rho(x,t) = q(x,t;x,t) = \int \int G(x_0,0;x,t)G(x,t;x_T,T)q(x_0,x_T)dx_0 \, dx_T. \tag{2.33}$$

In the Markov case, this reduces to

$$\rho(x,t) = q_f(x,t) q_b(x,t),$$  (2.34)

where

$$q_f(x,t) = \int G(x_0,0;x,t) q_f(x_0) dx_0,$$  (2.35a)

$$q_b(x,t) = \int G(x,t;x_T,T) q_b(x_T) dx_T$$  (2.35b)

satisfy, respectively, the forward and backward equations

$$\mathbf{L} q_f(x,t) = 0, \quad q_f(x,0) = q_f(x),$$  (2.36a)

$$\mathbf{L}^* q_b(x,t) = 0, \quad q_b(x,T) = q_b(x).$$  (2.36b)

The relations (2.34) and (2.36) illustrate the original idea of Schrödinger, which was to decompose the density $\rho(x,t)$ of the Markov process $x(t)$ into a product of forward and backward information "waves" $q_f(x,t)$ and $q_b(x,t)$, where the initial data for these waves is provided by the functions $q_f(x)$ and $q_b(x)$ obtained by solving the coupled integral equations (2.36).

The reciprocal and Markov diffusions introduced in Definition 2.1 can therefore be viewed as defined "in the sense of Schrödinger." Unlike Markov diffusions defined as the solutions of stochastic differential equations, which, for a given metric $\sigma_{ij}$, are specified by the $n$ components of the drift, and the initial density $\rho(x,0)$, the Markov diffusions of Definition 2.1 are parametrized by the $n+1$ potentials $\{A_i,\phi\}$, as well as the two end-point densities $q_f(x)$ and $q_b(x)$. This raises the question of whether these two types of Markov diffusions are equivalent. Such an equivalence was established by Jamison[4] and Nagasawa,[17] following an argument which is outlined below.

*Theorem 2.1*: $x(t)$ is a Markov diffusion in the sense of Definition 2.1 iff it is a Markov diffusion in the sense of stochastic differential equations.

*Proof*: Let $x(t)$ be a Markov diffusion in the sense of Definition 2.1. Define the gauge transformation

$$\chi_+(x,t) = \ln q_b(x,t).$$  (2.37)

Then, using the evolution equation (2.36b) for $q_b(x,t)$ inside the gauge relations (2.24a) and (2.24b), it is easy to verify that the transformed heat operator has the form

$$\mathbf{L}_+ = \frac{1}{2}\Delta - \nabla^i(b_{i+}) - \frac{\partial}{\partial t},$$  (2.38a)

$$b_{i+}(x,t) = -A_i(x,t) + \nabla_i \ln q_b(x,t),$$  (2.38b)

i.e., it is the forward operator corresponding to a stochastic differential equation with diffusion matrix $\sigma^{ij}(t)$ and drift $b^i_+$. Furthermore, using Eqs. (2.23) and (2.37), we find that for $t_0 = 0 < t_1 \cdots < t_p$, the joint density of $x(t_0)$, $x(t_1)$,..., $x(t_p)$ is given by

$$p(x_0,0;x_1,t_1;...;x_p,t_p) = \rho(x_0,0) \prod_{k=0}^{p-1} G_+(x_k,t_k;x_{k+1},t_{k+1}),$$  (2.39)

where $G_+$ is the Green's function of $\mathbf{L}_+$. This density is in the usual Markov form, so that $x(t)$ is a Markov diffusion in the usual sense.

Conversely, when $x(t)$ satisfies an Itô differential equation with forward operator $\mathbf{L}_+$ and initial density $\rho(x,0)$, by taking into account the normalization

$$Z_+(x,s;t) = \int G_+(x,s;y,t)dy \equiv 1 \tag{2.40}$$

inside the coupled integral equations

$$\rho(x_0,0) = q_f^+(x_0,0) \int G_+(x_0,0;x_T,T)q_b^+(x_T,T)dx_T , \tag{2.41a}$$

$$\rho(x_T,T) = q_b^+(x_T,T) \int G_+(x_0,0;x_T,T)q_f^+(x_0,0)dx_0 , \tag{2.41b}$$

it is easy to verify that these equations admit the solution

$$q_f^+(x,0) = \rho(x,0), \quad q_b^+(x,T) = 1. \tag{2.42a}$$

Furthermore, since the operator $\mathbf{L}_+$ does not contain a creation/killing term, the evolution equations (2.36) imply

$$q_f^+(x,t) = \rho(x,t), \quad q_b^+(x,t) = 1, \tag{2.42b}$$

so that for a Markov diffusion described by a forward stochastic differential equation, all the information about the diffusion process is contained in the forward wave, since the backward wave $q_b^+(x,t)$ is uniformly distributed. When identities (2.42a) are substituted inside the specification (2.39) of the joint probability densities of a Markov process, we find that these densities are in the form of Definition 2.1, so that Markov diffusions in the sense of stochastic differential equations satisfy Definition 2.1.                                                                    □

In the above proof, the gauge transformation (2.37) was employed to obtain a forward Itô differential equation for a Markov diffusion $x(t)$ defined in the sense of Definition 2.1. If instead of using Eq. (2.37), we select

$$\chi - (x,t) = -\ln q_f(x,t), \tag{2.43}$$

the heat operator is transformed into

$$\mathbf{L}_- = \frac{1}{2}\Delta - b_{i-}(x,t)\nabla^i - \frac{\partial}{\partial t}, \tag{2.44a}$$

$$b_{i-}(x,t) = -A_i(x,t) - \nabla_i \ln q_f(x,t), \tag{2.44b}$$

so that $x(t)$ satisfies also a backward stochastic differential equation with diffusion matrix $\sigma^{ij}(x)$ and drift $b^i_-$. In this case, the forward and backward information waves are given by

$$q_f^-(x,t) = 1, \quad q_b^-(x,t) = \rho(x,t), \tag{2.45}$$

so that all the information about the process $x(t)$ is now contained in the backward wave.

Finally, we verify that the reciprocal diffusions of Definition 2.1 satisfy the characterization of such diffusions given by Jamison.[4] According to this characterization, a process is a reciprocal diffusion iff it has the same three-point transition density $r(x,s;y,t;z,u)$ as a Markov

diffusion, but a different end-point density. But the expression (2.22) for the three-point transition density depends only on the Green's function $G$, i.e., on the potentials $\{A_i, \phi\}$. Thus, given a reciprocal diffusion $x(t)$, provided we replace its end-point density $q(x_0, x_T)$ by a factorable one, we obtain a Markov diffusion with the same three-point transition density, so that Jamison's criterion is satisfied.

To illustrate our construction procedure for Markov and reciprocal diffusions, we consider several examples.

*Example 2.1:* The *Wiener process* $W(t)$ is generated by the operator

$$\mathbf{L} = \frac{1}{2} \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial t} \qquad (2.46a)$$

with Green's function

$$G(x,s;y,t) = \frac{1}{(2\pi(t-s))^{1/2}} \exp - \frac{(y-x)^2}{2(t-s)} . \qquad (2.46b)$$

The marginal probability densities at the end points are

$$\rho(x,0) = \delta(x), \quad \rho(x,T) = \frac{1}{(2\pi T)^{1/2}} \exp - \frac{x^2}{2T}, \qquad (2.47)$$

so that

$$q_f(x,t) = \rho(x,t) = \frac{1}{(2\pi t)^{1/2}} \exp - \frac{x^2}{2t}, \quad q_b(x,t) = 1. \qquad (2.48)$$

*Example 2.2:* The *Brownian bridge* $B(t)$ can be expressed in terms of the Wiener process as

$$B(t) = W(t) - \frac{t}{T} W(T). \qquad (2.49)$$

It can be constructed from the heat operator (2.46a) and Green's function (2.46b), but with the marginal densities

$$\rho(x,0) = \rho(x,T) = \delta(x). \qquad (2.50)$$

In this case, the solution of the coupled integral equations (2.32a) and (2.32b) is given by

$$q_f(x) = q_b(x) = C\delta(x) \qquad (2.51)$$

with $C = (2\pi T)^{1/4}$, so that

$$q_f(x,t) = \frac{C}{(2\pi t)^{1/2}} \exp - \frac{x^2}{2t}, \qquad (2.52a)$$

$$q_b(x,t) = \frac{C}{(2\pi(T-t))^{1/2}} \exp - \frac{x^2}{2(T-t)}, \qquad (2.52b)$$

and

$$\rho(x,t) = \left(\frac{T}{2\pi t(T-t)}\right)^{1/2} \exp-\frac{x^2 T}{2t(T-t)} .$$  (2.53)

Note that the operator **L** is well-behaved, even though the drifts

$$b_+(x,t) = -\frac{x}{(T-t)} , \quad b_-(x,t) = \frac{x}{t}$$  (2.54)

of the forward and backward stochastic differential equations satisfied by $x(t)$ have singularities at $t=T$ and $t=0$, respectively, reflecting the fact that at these two times, the process density is totally concentrated at the node $x=0$.    □

*Example 2.3:* The *Bessel process* $R(t)$ is a Markov process defined as the radial part

$$R(t) = \left(\sum_{i=1}^{n} W_i^2(t)\right)^{1/2}$$  (2.55)

of a multivariate Wiener process whose components $W_i(t)$ are independent. It can be constructed[18] from the forward operator

$$\mathbf{L}_+ = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}\left(\frac{n-1}{2x} \cdot\right) - \frac{\partial}{\partial t}, \quad x \geqslant 0$$  (2.56a)

with Green's function

$$G_+(x,s;y,t) = \frac{\exp-(x^2+y^2)/(t-s)}{(t-s)(xy)^{n/2-1}} x^{n-1} I_{n/2-1}\left(\frac{xy}{t-s}\right),$$  (2.56b)

where $I_{n/2-1}$ is a modified Bessel function. Given the initial density

$$\rho(x,0) = \delta(x),$$  (2.57a)

we find that the density of $R(t)$ is given by

$$\rho(x,t) = \frac{2C_n x^{n-1}}{(2\pi t)^{n/2}} \exp-\frac{x^2}{2t},$$  (2.57b)

where $C_n$ is a normalizing constant. The drift $b^+(x,t) = (n-1)/2x$ corresponding to the above description is singular at $x=0$, reflecting the fact that the density $\rho(0,t)=0$ for all $t>0$, i.e., $x=0$ is a node of the diffusion $R(t)$ at all positive times. Similarly, if we consider the backward stochastic differential equation satisfied by $R(t)$, we find that the drift

$$b_-(x,t) = b_+(x,t) - \frac{\partial}{\partial x} \ln \rho(x,t) = \frac{x}{t} - \frac{n-1}{2x}$$  (2.58)

has also a singularity at $x=0$. On the other hand, the drift, i.e., the covector potential, can be completely eliminated by applying the gauge transformation

$$\chi(x,t) = -\frac{n-1}{2} \ln x$$  (2.59)

to the forward heat operator $\mathbf{L}_+$. In this case, the transformed heat operator takes the form

$$L = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{(n-1)(n-3)}{8x^2} - \frac{\partial}{\partial t}, \tag{2.60}$$

where all the information has been pushed into the scalar potential $\phi(x,t) = -(n-1)(n-3)/8x^2$. But this new description contains some information that did not appear explicitly in the forward description (2.56a). Specifically, we see that for $n=2$, the potential is repulsive, for $n=3$, it is identically zero, i.e., the particle is "free," and for $n>3$, the potential is attractive.□

The last example illustrates the flexibility afforded by the use of gauge transformations in the description of reciprocal and Markov diffusions.

## III. PATH INTEGRAL REPRESENTATION

In this section, we obtain a path integral representation of the Green's function $G(x,s;y,t)$ corresponding to the heat operator $L$. Although most applications of path integrals have been in quantum mechanics and field theory,[19] following Onsager and Machlup's original investigation,[20] path integral techniques have also been employed to assign a probability measure to the paths of diffusion processes. This line of work was originally formal, since it required the evaluation of an action functional along the nondifferentiable trajectories of a diffusion process, but it was subsequently given a rigorous meaning[21] when it was shown that the path measure can be interpreted as the probability that the diffusion stays within a tube of small radius centered about the path under consideration.

For our purposes, the most relevant work is that of Graham,[22] who obtained a path integral representation of the transition density of diffusion processes with creation and killing. Graham's result is more general than the one derived here, since it allows the metric to be space-dependent. However, the derivation of path integrals in curved space has been a subject of controversy, because when $\sigma_{ij}$ depends on $x$, the correspondence rule $p_j \leftrightarrow -\nabla_j$ which was used in Sec. II to construct $H$ from $H(x,p,t)$ cannot be defined unambiguously. Different choices of correspondence rule yield operators $H$ which differ by the addition of terms proportional to the scalar curvature of the space. The reader is referred to Ref. 23 for a discussion of path integrals in curved spaces.

The method which is employed here to derive a path integral representation of $G(x,s;y,t)$ is identical to Feynman's original approach. The Lagrangian $L(x,\dot{x},t)$ corresponding to the Hamiltonian (2.3) is given by

$$L(x,\dot{x},t) = \tfrac{1}{2}\sigma_{ij}(t)\dot{x}^i\dot{x}^j + A_i(x,t)\dot{x}^i - \phi(x,t), \tag{3.1}$$

and the action functional $\mathscr{S}(x_a,a;x_b,b)$ for a path $x(t)$ linking $x(a) = x_a$ to $x(b) = x_b$ with $0 \leqslant a \leqslant b \leqslant T$, can be expressed as

$$\mathscr{S}(x_a,a;x_b,b) = \int_a^b L(x(t),\dot{x}(t),t)dt. \tag{3.2}$$

Now, consider the discrete lattice $\mathbf{D}_h = \{t_k = a + kh; k=0,...,N\}$ obtained by dividing the interval $[a,b]$ into $N$ subintervals of length $h = (b-a)/N$. We seek to discretize the action functional $\mathscr{S}$ by replacing it by a discrete sum involving only the trajectory points $x(t_k)$ on the lattice $\mathbf{D}_h$. To achieve this objective, consider the colattice $\mathbf{D}'_h = \{a + (k + 1/2)h; k=0,...,N-1\}$. In the following, the mean and difference of successive discretization points are denoted as

$$m(t;h) = \tfrac{1}{2}(x(t+h/2) + x(t-h/2)), \tag{3.3a}$$

$$dx(t;h) = x(t+h/2) - x(t-h/2), \tag{3.3b}$$

with $t \in \mathbf{D}'_h$. Then, consider a path originating at $x(a)=x_a$, terminating at $x(b)=x_b$ and constituted of piecewise linear segments linking the points $x(t_k)$, with $t_k \in \mathbf{D}_h$. Over each segment $t-h/2 \leqslant s \leqslant t+h/2$, with $t \in \mathbf{D}'_h$, the motion of the particle is specified by

$$\ddot{x}_i(s)=0, \tag{3.4}$$

which reflects that in the absence of potentials, no forces act on the particle. This implies that over each segment $\dot{x}_i(s)=u_i$, i.e., the velocity, when expressed in contravariant form, is constant. In fact, it is easy to check that

$$u_i=\sigma_{ij}(t)\,\frac{dx^j}{h}+O(h^2). \tag{3.5}$$

For such a piecewise linear trajectory, the action functional $\mathscr{S}$ can be decomposed as

$$S(x_a,a;x_b,b)=\sum_{t\in\mathbf{D}'_h}\int_{t-h/2}^{t+h/2} L(x(s),\dot{x}(s),s)ds. \tag{3.6}$$

Performing a Taylor series expansion of $x(s)$, $\dot{x}(s)$ and $s$ in the vicinity of $m(t;h)$, $dx(t;h)/h$ and $t$, respectively, and taking into account (3.5), we find

$$\int_{t-h/2}^{t+h/2} L(x(s),\dot{x}(s),s)ds=L_h(m,dx,t)+O((dx)^3)+O(dx\,h)+O(h^2), \tag{3.7}$$

with

$$L_h(m,dx,t)=(1/2h)\sigma_{ij}(t)dx^i\,dx^j+A_i(m,t)dx^i-\phi(m,t)h. \tag{3.8}$$

The orders of truncation employed in Eq. (3.7) ensure that for all piecewise linear paths such that $dx(t;h)$ has size $h^{1/2}$, which is the standard deviation of the increments of a Markov diffusion, the action functional satisfies

$$S(x_a,a;x_b,b)=\sum_{t\in\mathbf{D}'_h} L_h(m,dx,t)+O(h^{1/2}). \tag{3.9}$$

Consider now the preexponential factor

$$C_h(t)=\frac{\sigma^{1/2}(t)}{(2\pi h)^{n/2}}, \tag{3.10}$$

where $\sigma(t)=\det\sigma_{ij}(t)$. We can assign to a piecewise linear path $x_h$ linking the points $x(a)=x_a$ and $x(b)=x_b$ and passing through the points $x(t_k)$ with $t_k\in\mathbf{D}_h$, the probability

$$p(x_h)=\prod_{t\in\mathbf{D}'_h} C_h(t)\exp-L_h(m,dx,t), \tag{3.11}$$

where the arguments of $m(t;h)$ and $dx(t;h)$ have been suppressed. Note that the path probabilities defined above are *unnormalized*, since as will be shown below, the sum of the probabilities for all piecewise linear paths linking $x(a)=x_a$ and $x(b)=x_b$ is not equal to 1. Also, we cannot claim at this stage that the probability assignment (3.11) is consistent with the probabilistic description given in Definition 2.1 for reciprocal or Markov diffusions constructed from the operator $\mathbf{L}$. The equivalence of these two descriptions is based on the following result.

**Theorem 3.1**: The Green's function $G(x_a, a; x_b, b)$ specified by Eq. (2.7) can be expressed as the path integral

$$G(x_a, a; x_b, b) = \lim_{h \to 0} \int \prod_{\tau \in \mathbf{D}_h} dx(\tau) \prod_{t \in \mathbf{D}_h'} C_h(t) \exp{-L_h(m, dx, t)}. \tag{3.12}$$

If $K(x_a, a; x_b, b)$ denotes the expression appearing on the right-hand side of Eq. (3.12), we see that it is obtained by summing the probabilities of all piecewise linear paths linking $x_a$ and $x_b$ and then letting the lattice spacing $h$ tend to zero. It is therefore a Feynman path integral, except that the path probabilities are real, instead of being complex probability amplitudes.

**Proof:** Let $f(.)$ be an arbitary function defined over $\mathbf{R}^n$, and define

$$f(y, t) \triangleq \int K(x, s; y, t) f(x) dx. \tag{3.13}$$

To prove Theorem 3.1, we only need to show that $f(y, t)$ satisfies the heat equation

$$\mathbf{L} f(y, t) = 0, \quad t \geqslant s \tag{3.14}$$

with initial condition $f(y, s) = f(y)$.

From Eq. (3.13), we have

$$f\left(y, t + \frac{h}{2}\right) = \int C_h(t) \exp(-L_h(m, z, t)) f\left(x, t - \frac{h}{2}\right) dx, \tag{3.15}$$

where $z = y - x$ and $m = (y + x)/2$. Replacing the integral over $x$ by an integral over $z$ yields

$$f\left(y, t + \frac{h}{2}\right) = \int C_h(t) \exp(-L_h(y - z/2, z, t)) f\left(y - z, t - \frac{h}{2}\right) dz, \tag{3.16}$$

where we can expand all terms inside the integral in the vicinity of $y$ and $t$. This gives

$$C_h(t) \exp{-L_h(y - z/2, z, t)} = M_h(y, z, t) \exp{-E_h(y, z, t)} \tag{3.17}$$

with

$$E_h(y, z, t) = \frac{\sigma_{ij}(t)}{2h} (z^i + A^i(y, t) h)(z^j + A^j(y, t) h), \tag{3.18a}$$

$$M_h(y, z, t) = \frac{\sigma^{1/2}(t)}{(2\pi h)^{n/2}} \left[ 1 + \left( \frac{1}{2} A^i(y, t) A_i(y, t) + \phi(y, t) \right) h + \frac{1}{2} \frac{\partial A_i}{\partial y^k} (y, t) z^i z^k + O(z^3) + O(zh) \right. $$
$$\left. + O(h^2) \right], \tag{3.18b}$$

where the exponent $E_h$ is quadratic in $z$. For small $z$, we have also the approximation

$$f\left(y - z, t - \frac{h}{2}\right) = f(y, t) - \frac{\partial f}{\partial y^k} (y, t) z^k - \frac{1}{2} \frac{\partial f}{\partial t} (y, t) h + \frac{1}{2} \frac{\partial^2 f}{\partial y^k \partial y^l} (y, t) z^k z^l + O(z^3) + O(zh) + O(h^2). \tag{3.19}$$

Using standard identities for the first- and second-order moments of a Gaussian distribution, and identifying the terms proportional to $h$ on both sides of Eq. (3.16) yields

$$\frac{\partial f}{\partial t}(y,t) = \left(\frac{A^i A_i}{2} + \phi + \frac{\sigma^{ij}}{2}\frac{\partial A_i}{\partial y^j}\right)(y,t) + A^i(y,t)\frac{\partial f}{\partial y^i}(y,t) + \frac{\sigma^{ij}}{2}\frac{\partial^2 f}{\partial y^i \partial y^j}(y,t), \qquad (3.20)$$

so that $f(y,t)$ satisfies the heat equation (3.14). Furthermore, if we consider the generator

$$G_h(x,t-h/2;y,t+h/2) = C_h(t)\exp - L_h(m,z,t), \qquad (3.21)$$

with $m=(y+x)/2$ and $z=y-x$, the expression (3.8) for $L_h(m,z,t)$ implies

$$\lim_{h\to 0} G_h(x,s;y,s+h) = \delta(y-x), \qquad (3.22)$$

which ensures that $f$ obeys the initial condition $f(y,s)=f(y)$.       $\square$

The generator $G_h(x,t-h/2;y,t+h/2)$ given by Eq. (3.21) is consistent with the short-time asymptotic expansions of the transition density of Markov diffusions obtained in Ref. 24 for the more general case of a space-dependent metric.

Theorem 3.1 and the preceding analysis yield the local short-time asymptotic expansion

$$G(x,s;y,s+h) = G_h(x,s;y,s+h)[1+O(h^{3/2})] \qquad (3.23)$$

for $y-x=O(h^{1/2})$, which plays an essential role in characterizing the local behavior of reciprocal diffusions.

## IV. STOCHASTIC NEWTON LAW

In this section, we derive Newton's law of reciprocal diffusions. Our derivation is patterned after that of Ref. 8, which used a short-time asymptotic expansion for the transition density of Markov diffusions to approximate locally the three-point transition density of reciprocal diffusions. The only difference is that we employ here the expansion (3.23) for $G$. By substituting the short-time asymptotic expansion (3.23) inside the conditional density $r(x(t-h),t-h;x(t),t;x(t+h),t+h)$ of $x(t)$ given $x(t-h)$ and $x(t+h)$, with $h$ small, we find it admits the approximation

$$r(x(t+h),t+h;x(t),t;x(t-h),t-h)\approx M\exp - E \qquad (4.1)$$

with

$$E = L_h(m_+,dx_+,t_+) + L_h(m_-,dx_-,t_-), \qquad (4.2)$$

where

$$t_\pm = t \pm h/2, \quad m_\pm = m(t\pm h/2;h), \quad dx_\pm = dx(t\pm h/2;h), \qquad (4.3)$$

and $M = M(x(t+h),x(t-h),t;h)$ is a constant independent of $x(t)$. Our derivation of Newton's law relies on a local approximation of $r(x(t+h),t+h;x(t),t;x(t-h),t-h)$ for small $h$.

Consider the centered mean and centered first- and second-order differences

$$\bar{x}(t;h) = \tfrac{1}{2}(x(t+h)+x(t-h)), \qquad (4.4a)$$

$$d^1 x(t;h) = \tfrac{1}{2}(x(t+h)-x(t-h)), \qquad (4.4b)$$

$$d^2x(t;h) = x(t+h) + x(t-h) - 2x(t), \tag{4.4c}$$

so that

$$x(t \pm h) = \bar{x}(t;h) \pm d^1x(t;h), \tag{4.5a}$$

$$x(t) = \bar{x}(t;h) - \tfrac{1}{2}d^2x(t;h). \tag{4.5b}$$

Then, the means $m_\pm$ and differences $dx_\pm$ given by Eq. (4.3) can be expressed as

$$m_\pm = \bar{x}(t;h) + \tfrac{1}{2}(\pm d^1x(t;h) - \tfrac{1}{2}d^2x(t;h)), \tag{4.6a}$$

$$dx_\pm = d^1x(t;h) \pm \tfrac{1}{2}d^2x(t;h). \tag{4.6b}$$

Expanding $m_\pm$ in the vicinity of $\bar{x}(t;h)$ and $t_\pm$ in the vicinity of $t$ inside the discretized Lagrangian $L_h(m_\pm, dx_\pm, t_\pm)$ and substituting inside $E$ in (4.2) yields a local approximation of $r(x(t+h),t+h;x(t),t;x(t-h),t-h)$, the conditional density of $x(t)$ given $x(t+h)$ and $x(t-h)$, for small $h$. In deriving this approximation, we employ the following rules.

(a) Since we are only interested in the local distribution of $x(t)$, or more precisely $d^2x(t;h)$, all the terms that do not contain $d^2x(t;h)$ can be ignored, since they can always be absorbed inside the constant $M$ of Eq. (4.1).

(b) To decide which terms to keep, and which ones to neglect in our expansions, we assume that $d^1x(t;h)$ has size $h^{1/2}$, which is the standard deviation of the increments of a Markov diffusion. In approximating $L_h$, we neglect all terms proportional to $O((d^2x)^{2p} \times h^{(3/2-p)})$ for $p \geqslant 1$, and $O((d^2x)^{(2p+1)}h^{(1-p)})$ for $p \geqslant 0$, with $p$ integer. The motivation for selecting these orders of truncation is that we want to evaluate accurately all the moments of the conditional density of $d^2x(t;h)$ given $d^1x(t;h)$ and $\bar{x}(t;h)$ up to order $h^2$.

We obtain

$$L_h(m_+, dx_+, t_+) \approx \frac{1}{2h}\left[\sigma_{ij}(t) + \frac{1}{2}\frac{\partial \sigma_{ij}}{\partial t}(t)h\right]\left(d^1x + \frac{d^2x}{2}\right)^i\left(d^1x + \frac{d^2x}{2}\right)^j$$
$$+ \left[A_i(\bar{x},t) + \frac{1}{2}\frac{\partial A_i}{\partial x^j}(\bar{x},t)\left(d^1x - \frac{d^2x}{2}\right)^j + \frac{1}{2}\frac{\partial A_i}{\partial t}(\bar{x},t)h\right]\left(d^1x + \frac{d^2x}{2}\right)^i$$
$$- \left[\phi(\bar{x},t) + \frac{1}{2}\frac{\partial \phi}{\partial x^i}(\bar{x},t)\left(d^1x - \frac{d^2x}{2}\right)^i\right]h, \tag{4.7a}$$

$$L_h(m_-, dx_-, t_-) \approx \frac{1}{2h}\left[\sigma_{ij}(t) - \frac{1}{2}\frac{\partial \sigma_{ij}}{\partial t}(t)h\right]\left(d^1x - \frac{d^2x}{2}\right)^i\left(d^1x - \frac{d^2x}{2}\right)^j$$
$$+ \left[A_i(\bar{x},t) - \frac{1}{2}\frac{\partial A_i}{\partial x^j}(\bar{x},t)\left(d^1x + \frac{d^2x}{2}\right)^j - \frac{1}{2}\frac{\partial A_i}{\partial t}(\bar{x},t)h\right]\left(d^1x - \frac{d^2x}{2}\right)^i$$
$$- \left[\phi(\bar{x},t) - \frac{1}{2}\frac{\partial \phi}{\partial x^i}(\bar{x},t)\left(d^1x + \frac{d^2x}{2}\right)^i\right]h. \tag{4.7b}$$

This yields

$$E \approx E_0 + \frac{\sigma_{ij}(t)}{4h} d^2x^i d^2x^j + \frac{1}{2} \left( \frac{\partial \sigma_{ij}}{\partial t} - dA_{ij} \right) (\bar{x},t) d^2x^i d^1x^j + \left( \frac{\partial \phi}{\partial x^i} + \frac{\partial A_i}{\partial t} \right) (\bar{x},t) d^2x^i h, \quad (4.8)$$

where the term $E_0$ does not contain $d^2x$ and

$$dA_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} \qquad (4.9)$$

denotes the exterior derivative of $A_i$.

The expression (4.8) for $E$ indicates that the three-point transition density $r(x(t+h),t +h;x(t),t;x(t-h),t-h)$ can be approximated locally by a Gaussian distribution for $d^2x(t;h)$. The conditional mean of $d^2x(t;h)$ given $x(t\pm h)$ is obtained by minimizing $E$ with respect to $d^2x^i$. Dividing both sides of the resulting expression by $h^2$, and denoting by $u^j(t;h) = d^1x^j(t;h)/h$ the empirical velocity, we obtain the *stochastic Newton law*

$$a_i = F_i(\bar{x},u,t) + O(h^{1/2}), \qquad (4.10)$$

where

$$a_i = E \left[ \sigma_{ij}(t) \frac{d^2x^j}{h^2} + \frac{\partial \sigma_{ij}}{\partial t} (t) u^j \, | \, x(t\pm h) = \bar{x} \pm uh \right] \qquad (4.11)$$

is the mean acceleration in contravariant form, and

$$F_i(\bar{x},u,t) = dA_{ij}(\bar{x},t) u^j - \left( \frac{\partial \phi}{\partial x^i} + \frac{\partial A_i}{\partial t} \right) (\bar{x},t) \qquad (4.12)$$

is the force applied to a particle with position $\bar{x}$ and empirical velocity $u$ at time $h$. By analogy with the Lorentz force law of electromagnetism, the first term in Eq. (4.12) can be viewed as a magnetic force (because of the skew symmetry of $dA_{ij}$, this force is orthogonal to the velocity vector $u^i$), and the second term as an electric force.

Since the coefficient of $d^2x^i d^2x^j/2$ in $E$ is $\sigma_{ij}(t)/2h$, we can conclude that the conditional covariance matrix of $d^2x(t;h)$ given $x(t\pm h)$ is given by

$$E[d^2x^i d^2x^j \, | \, x(t\pm h) = \bar{x} \pm uh] = 2\sigma^{ij}(t)h + O(h^{5/2}). \qquad (4.13)$$

Note that the evaluation of the size of the error terms in (4.10) and (4.13) relies on the assumption that $d^1x$ has size $O(h^{1/2})$. We have therefore proved the following result.

**Theorem 4.1**: If $x(t)$ is a reciprocal diffusion constructed from the triple $\{A_i(x,t),\phi(x,t),q(x_0,x_T)\}$, it satisfies the stochastic Newton law (4.10)–(4.13).

An important aspect of Newton's law is that it is *noncausal*, in the sense that it specifies the conditional acceleration $d^2x(t;h)/h^2$ at time $t$ given the positions of the diffusing particle at times $t-h$ and $t+h$, so that the knowledge of both past and future is necessary to specify the motion at a fixed time. From a physical point of view, the local motion of a diffusing particle is obtained by adding to the classical motion specified by the forces $F_i(x,u,t)$ some random fluctuations, which according to (4.13) have a size proportional to $h^{1/2}$. The statistical properties of the random fluctuations are not yet well understood, except in the Gaussian case, for which it was shown[25] that the fluctuations are locally correlated in time, and such that the correlation between two successive time increments depends on the forces acting on the diffusing particle.

It is worth noting that the stochastic Newton law derived here differs from the one proposed by Nelson,[5] since it relies on a different definition of the stochastic acceleration. According to Nelson, the stochastic acceleration of a Markov diffusion $x(t)$ is given by

$$a_N^i(t) = \tfrac{1}{2}(D_- D_+ + D_+ D_-)x^i(t),\tag{4.14}$$

where $D_+$ and $D_-$ denote the mean forward and backward derivatives corresponding to $x(t)$. In flat space, for an arbitrary tensor $T(x,t)$, these derivatives are defined as

$$D_+ T(x(t),t) = \lim_{h \to 0} \frac{1}{h} E_t[T(x(t+h),t+h) - T(x(t),t)],\tag{4.15a}$$

$$D_- T(x(t),t) = \lim_{h \to 0} \frac{1}{h} E_t[T(x(t),t) - T(x(t-h),t-h)],\tag{4.15b}$$

where $E_t$ denotes the conditional expectation given $x(t)$. To verify that definitions (4.11) and (4.14) yield different values for the acceleration, let $x(t)$ be a Markov diffusion with potentials $\{A_i,\phi\}$. Then, if $b_{i+}$ and $b_{i-}$ are the forward and backward drifts given by Eqs. (2.38b) and (2.44b), respectively, and if we introduce the *mean velocity*

$$v_i(x,t) = \tfrac{1}{2}(b_{i+}(x,t) + b_{i-}(x,t)),\tag{4.16}$$

Nelson's version of Newton's law takes the form

$$a_{Ni}(t) = F_{Ni}(x(t),v(x(t),t),t),\tag{4.17}$$

where the forces $F_{Ni}$ are obtained[11] by replacing the potentials $\{A_i,\phi\}$ by

$$A_{Ni} = A_i,\tag{4.18a}$$

$$\phi_N = \phi + \frac{\Delta\rho^{1/2}}{\rho^{1/2}}\tag{4.18b}$$

inside Eq. (4.12). Thus, in Nelson's formulation the potentials used to contruct a Markov diffusion do not match those appearing in the force term of its Newton law. Furthermore, the correction term which is added to the scalar potential $\phi$ to obtain $\phi_N$ is nonlocal in the sense that a change in the probability density $\rho(x,t)$ at one point in space instantaneously affects the potentials at points far away from it. This instantaneous action at a distance is rather objectionable since no such effect can exist in classical physics. In contrast, the version of Newton's law given in Theorem 4.1 is completely free of nonlocal effects.

As indicated in Sec. II, the potentials $\{A_i(x,t),\phi(x,t)\}$ corresponding to a reciprocal diffusion $x(t)$ are only specified modulo a gauge transformation of the form (2.24a) and (2.24b). Obviously, the choice of gauge does not affect the forces appearing in Newton's law. In this context, the concept of local reciprocal invariants introduced by Clark[9] can be formulated as follows.

*Definition 4.1*: Two reciprocal diffusions in a space with metric $\sigma_{ij}(t)$ are locally equivalent if for identical positions $x(t \pm h) = \bar{x} \pm uh$, they are subjected to the same forces. Quantities that are preserved under local equivalence are said to be local invariants. A set of local invariants is complete if two diffusions with identical invariants are locally equivalent.

If two reciprocal diffusions with potentials $\{A_i,\phi\}$ and $\{A_i',\phi'\}$ are locally equivalent, we see from (4.12) that, because the position $\bar{x}$ and velocity $u$ are arbitrary, we must have

$$dA_{ij} = dA'_{ij}, \tag{4.19a}$$

$$\frac{\partial \phi}{\partial x^i} + \frac{\partial A_i}{\partial t} = \frac{\partial \phi'}{\partial x^i} + \frac{\partial A'_i}{\partial t}, \tag{4.19b}$$

from which we can deduce that the potentials $\{A_i, \phi\}$ and $\{A'_i, \phi'\}$ must be related by a gauge transformation. Thus, we have proved the following result, due to Clark.[9]

**Theorem 4.2**: Two reciprocal diffusions are locally equivalent if and only if their potentials are related by a gauge transformation. Furthermore, the $n(n+1)/2$ quantities

$$\left\{ dA_{ij}, \frac{\partial \phi}{\partial x^i} + \frac{\partial A_i}{\partial t} \right\} \tag{4.20}$$

constitute a complete set of local invariants.

Theorem 4.2 implies that if two reciprocal diffusions are locally equivalent, they have the same three-point transition densities, but the end-point density $q(x_0, x_T)$ may be different. Thus, the class of equivalence of a given reciprocal diffusion contains the Markov diffusions obtained by replacing $q(x_0, x_T)$ by a factorable density of the form $q_f(x_0) q_b(x_T)$.

## V. GIRSANOV TRANSFORMATION

Consider two reciprocal diffusions constructed from the triples $\{A_i(x,t), \phi(x,t), q(x_0, x_T)\}$ and $\{A'_i(x,t), \phi'(x,t), q'(x_0, x_T)\}$, defined over the interval $[0, T]$, in a space with metric $\sigma_{ij}(t)$. When these two diffusions are discretized over the lattice $\mathbf{D}_h = \{t_k = kh; \ k = 0, ..., N\}$ with $h = T/N$, the joint densities of the discretized processes can be approximated as

$$p(x_0, 0; ...; x_k, t_k; ...; x_N, T) = \left[ q(x_0, x_N) \prod_{k=0}^{N-1} G_h(x_k, t_k; x_{k+1}, t_{k+1}) \right] (1 + O(h^{1/2})), \tag{5.1a}$$

$$p'(x_0, 0; ...; x_k, t_k; ...; x_N, T) = \left[ q'(x_0, x_N) \prod_{k=0}^{N-1} G'_h(x_k, t_k; x_{k+1}, t_{k+1}) \right] (1 + O(h^{1/2})) \tag{5.1b}$$

with

$$G_h(x(t+h/2), t+h/2; x(t-h/2), t-h/2) = C_h(t) \exp - L_h(m, dx, t), \tag{5.2a}$$

$$G'_h(x(t+h/2), t+h/2; x(t-h/2), t-h/2) = C'_h(t) \exp - L'_h(m, dx, t), \tag{5.2b}$$

where the arguments of $m(t; h)$ and $dx(t; h)$ have been suppressed, and we have assumed $dx = O(h^{1/2})$. But the preexponential factor $C_h(t)$ depends only on the metric $\sigma_{ij}$, and is the same for both processes. Consequently, the likelihood-ratio function for the two discretized processes can be approximated as

$$\Lambda(\mathbf{x}_h) \overset{\Delta}{=} \frac{p'(x_0, 0; ...; x_k, t_k; ...; x_N, T)}{p(x_0, 0; ...; x_k, kh; ...; x_N, T)} = \left( \frac{q'(x_0, x_N)}{q(x_0, x_N)} \prod_{t \in \mathbf{D}'_h} \exp(L_h - L'_h)(m, dx, t) \right) (1 + O(h^{1/2})), \tag{5.3}$$

with

$$(L_h - L'_h)(m, dx, t) = (A_i - A'_i)(m, t) dx^i - (\phi - \phi')(m, t) h. \tag{5.4}$$

The likelihood-ratio function $\Lambda(\mathbf{x}_h)$ can be interpreted as the ratio of the probabilities of a piecewise linear path passing through the point $x(t_k)$, $0 \leqslant k \leqslant N$ under the probability measures $\mathbf{P}'$ and $\mathbf{P}$ corresponding to $\{A'_i, \phi', q'\}$ and $\{A_i, \phi, q\}$, respectively. Letting $h \to 0$ in Eq. (5.3), we find that the likelihood-ratio, or Radon–Nikodym derivative of $\mathbf{P}'$ with respect to $\mathbf{P}$ is given by

$$\Lambda = \frac{d\mathbf{P}'}{d\mathbf{P}} = \frac{q'(x_0, x_T)}{q(x_0, x_T)} \exp \delta S(x_T, T; x_0, 0),$$   (5.5)

where the action difference $\delta S$ is defined as

$$\delta S(x_T, T; x_0, 0) \triangleq \lim_{h \to 0} \sum_{\mathbf{D}'_h} (L_h - L'_h)(m, dx, t) = \int_0^T (A_i - A'_i)(x, t) dx^i - \int_0^T (\phi - \phi')(x, t) dt,$$   (5.6)

with the limit being taken in the mean-square sense. Both integrals in Eq. (5.6) concern a sample path $x(t)$ linking $(x_0, 0)$ and $(x_T, T)$. The first integral is defined in the sense of Stratonovich, whereas the second integral is a standard Riemann integral. Note that $\Lambda$ is well defined only if the end-point density $q'(x_0, x_T)$ is absolutely continuous with respect to $q(x_0, x_T)$. By analogy with the interpretation for discretized paths, $\Lambda$ can be viewed as the ratio of the probabilities of a path $x(t)$ under $\mathbf{P}'$ and $\mathbf{P}$. Although a given path has zero probability under each of these measures, the ratio (5.5) is well defined. Thus, we have proved the following generalization of the Girsanov transformation of Markov diffusions.

**Theorem 5.1**: The likelihood ratio function $\Lambda$ of two reciprocal diffusions $\{A'_i, \phi', q'\}$ and $\{A_i, \phi, q\}$, where $q'$ is absolutely continuous with respect to $q$ is given by Eqs. (5.5) and (5.6).

For the case of Markov diffusions, we have

$$q'(x_0, x_T) = q'_f(x_0) q'_b(x_T), \quad q(x_0, x_T) = q_f(x_0) q_b(x_T).$$   (5.7)

Furthermore, by an appropriate choice of gauge transformations, these two diffusions can be represented entirely in terms of forward propagating information waves, i.e.,

$$q'_f(x, t) = \rho'(x, t), \quad q'_b(x, t) \equiv 1,$$   (5.8a)

$$q_f(x, t) = \rho(x, t), \quad q_b(x, t) \equiv 1,$$   (5.8b)

where the two Markov diffusions satisfy forward stochastic differential equations with drifts $b''(x, t)$ and $b^i(x, t)$, respectively. By making the identification [see Eq. (2.14)]

$$A'_i = -b'_i, \quad \phi' = -\tfrac{1}{2}(b'^i b'_i + \nabla^i b'_i),$$   (5.9a)

$$A_i = -b_i, \quad \phi = -\tfrac{1}{2}(b^i b_i + \nabla^i b_i),$$   (5.9b)

it is easy to verify that Eqs. (5.5) and (5.6) reduce to

$$\Lambda = \frac{\rho'(x_0, 0)}{\rho(x_0, 0)} \exp \delta S(x_T, T; x_0, 0)$$   (5.10)

with

$$\delta S(x_T, T; x_0, 0) = \int_0^T (b'_i - b_i)(x, t) dx^i - \frac{1}{2} \int_0^T [(b'^i b'_i + \nabla^i b'_i) - (b^i b_i + \nabla^i b_i)] dt,$$   (5.11)

which is the standard expression for the Girsanov transformation of Markov diffusions.

Note that the expressions (5.5) and (5.6) for the Girsanov transformation of reciprocal diffusions are simpler than the expression obtained by Clark,[9] which required a double integral. However, Clark's integral was expressed entirely in terms of local reciprocal invariants, whereas the path integral appearing in the expression (5.6) for $\delta S(x_T,T;x_0,0)$ is not gauge invariant. Specifically, if we consider the gauge transformation

$$\tilde{A}_i(x,t) = A_i(x,t) - \frac{\partial \chi}{\partial x^i}(x,t),$$            (5.12a)

$$\tilde{\phi}(x,t) = \phi(x,t) + \frac{\partial \chi}{\partial t}(x,t),$$            (5.12b)

we have the path integral transformation rule

$$\int_0^T \tilde{A}_i \, dx^i - \int_0^T \tilde{\phi} \, dt = \left( \int_0^T A_i \, dx^i - \int_0^T \phi \, dt \right) - (\chi(x_T,T) - \chi(x_0,0)).$$            (5.13)

The value of the integral is changed by the quantity $\chi(x_T,T) - \chi(x_0,0)$ which is *path independent*, and depends only on the end points of the path. However, according to Eq. (2.23b), under the gauge transformation (5.12), the end-point density is transformed according to the rule

$$\tilde{q}(x_0,x_T) = \exp(\chi(x_T,T) - \chi(x_0,0))q(x_0,x_T),$$            (5.14)

which removes the additional term introduced by Eq. (5.13). Thus, the likelihood-ratio function $\Lambda$ is gauge invariant.

## VI. VELOCITY DISTRIBUTION

The Newton law (4.10)–(4.13) captures the local structure of the conditional density of $x(t)$ given $x(t-h)$ and $x(t+h)$, or equivalently, given $\bar{x}(t;h)$ and $d^1x(t;h)$, for small $h$. To obtain a full probabilistic description of $x(t)$, we must also characterize the joint density of $x(t-h)$ and $x(t+h)$ for small $h$. This will be accomplished by finding a local approximation of the conditional density of $d^1x(t;h)$ given the mean position $\bar{x}(t;h)$. Specifically, it will be shown that this conditional density is approximately Gaussian, and when expressed in terms of the empirical velocity $u(t;h) = d^1x(t;h)/h$, it can be viewed as specifying a velocity distribution, which except for the fact it applies to a diffusion, is reminiscent of the Maxwell velocity distribution for a dilute gas locally in thermal equilibrium.

Let $p(x,t-h;y,t+h)$ be the joint probability density of $x(t-h)$ and $x(t+h)$. By taking into account the expression (2.31) relating the two-point density to the Green's function $G$, and the short-time asymptotic expansion (3.23), we find it admits the local approximation

$$p(x,t-h;y,t+h) = q(x,t-h;y,t+h)C_{2h}(t)\exp - L_{2h}(\bar{x},2d^1x,t) + O(h^{3/2}),$$            (6.1)

with $\bar{x} = (x+y)/2$, $d^1x = (y-x)/2$, where it is assumed that $d^1x = O(h^{1/2})$. Let

$$\theta(x,s;y,t) \triangleq \ln q(x,s;y,t).$$            (6.2)

Expanding $y = \bar{x} + d^1x$ and $x = \bar{x} - d^1x$ in the vicinity of $\bar{x}$, and $t \pm h$ in the vicinity of $t$, inside $\theta(y,t-h;x,t+h)$ yields

$$\theta(x,t-h;y,t+h) = \theta(\bar{x},t;\bar{x},t) + 2w_i(\bar{x},t)d^1x^i + 2\pi_{ij}(\bar{x},t)d^1x^i + O(h^{3/2}), \qquad (6.3)$$

with

$$w_i(\bar{x},t) \triangleq \frac{1}{2}\left(\frac{\partial\theta}{\partial y^i} - \frac{\partial\theta}{\partial x^i}\right)(\bar{x},t;\bar{x},t), \qquad (6.4a)$$

$$\pi_{ij}(\bar{x},t) \triangleq \frac{1}{4}\left[\left(\frac{\partial}{\partial y^i} - \frac{\partial}{\partial x^i}\right)\left(\frac{\partial}{\partial y^j} - \frac{\partial}{\partial x^j}\right)\theta\right](\bar{x},t;\bar{x},t). \qquad (6.4b)$$

Noting that $\rho(x,t) = q(x,t;x,t)$, this gives the approximation

$$q(x,t-h;y,t+h) = \rho(\bar{x},t)\exp(2w_i d^1x^i + 2\pi_{ij}d^1x^i\,d^1x^j) + O(h^{3/2}). \qquad (6.5)$$

In the Markov case, by observing that

$$q(x,s;y,t) = q_f(x,s)q_b(y,t), \qquad (6.6)$$

where $q_f(x,t)$ and $q_b(x,t)$ are given by Eqs. (2.35a) and (2.35b), we obtain

$$\theta(x,s;y,t) = \theta_f(x,s) + \theta_b(y,t) \qquad (6.7)$$

with

$$\theta_f(x,t) = \ln q_f(x,t), \quad \theta_b(x,t) = \ln q_b(x,t). \qquad (6.8)$$

In this case, the expressions (6.4a) and (6.4b) reduce to

$$w_i(x,t) = \frac{1}{2}(\nabla_i \ln q_b(x,t) - \nabla_i \ln q_f(x,t)) = \nabla_i \ln\left(\frac{q_b(x,t)}{q_f(x,t)}\right)^{1/2}, \qquad (6.9)$$

$$\pi_{ij}(x,t) = \tfrac{1}{4}\nabla_i\nabla_j(\ln q_f(x,t) + \ln q_b(x,t)) = \tfrac{1}{4}\nabla_i\nabla_j \ln \rho(x,t). \qquad (6.10)$$

This shows that in the Markov case, $w^i(x,t)$ is a gradient function, and the tensor $\pi_{ij}(x,t)$ is proportional to the Hessian of the log density $\ln \rho(x,t)$.

From Eq. (3.8), we also see that

$$L_{2h}(\bar{x},2d^1x,t) = \sigma_{ij}(t)d^1x^i\,d^1x^j/h - 2A_i(\bar{x},t)d^1x^i + O(h). \qquad (6.11)$$

We are now ready to find an approximation of the joint density $p(x,t-h;y,t+h)$ of $x(t-h)$ and $x(t+h)$ for small $h$. However, instead of approximating $p(x,t-h;y,t+h)$, it is more convenient to approximate the joint density $p(\bar{x},d^1x,t;h)$ of $\bar{x}(t;h)$ and the difference $d^1x(t;h)$. Noting that the Jacobian of the transformation from $x(t-h)$, $x(t+h)$ to $\bar{x}(t;h)$ and $d^1x(t;h)$ is $2^n$, and substituting approximations (6.5) and (6.11) inside (6.1) yields

$$p(\bar{x},d^1x,t;h) = 2^n p(y,t-h;x,t+h) = \rho(\bar{x},t)\,\frac{\sigma^{1/2}(t)}{(\pi h)^{n/2}}$$

$$\times\exp\left[-\left(\frac{\sigma_{ij}}{h} - 2\pi_{ij}\right)d^1x^i\,d^1x^j + 2(w_i - A_i)d^1x^i\right](1 + O(h)). \qquad (6.12)$$

From Eq. (6.12), we see that for small $h$ the density $p(\bar{x},t;h)$ of $\bar{x}(t;h)$ admits the approximation

$$p(\bar{x},t;h) = \rho(\bar{x},t) + O(h).$$                    (6.13)

Also, the conditional density of $d^1x(t;h)$ given $\bar{x}(t;h)$ can be approximated locally by a Gaussian distribution with mean

$$E[d^1x^i(t;h) | \bar{x}(t;h) = \bar{x}] = v^i(\bar{x},t)h + O(h^2),$$                    (6.14)

where the *mean velocity* is given by

$$v^i(x,t) \triangleq -A^i(x,t) + w^i(x,t).$$                    (6.15)

The conditional covariance of $d^1x(t;h)$ is obtained by inverting the tensor

$$\gamma_{ij} = (2/h)(\sigma_{ij}(t) - 2\pi_{ij}(x,t)h).$$                    (6.16)

This gives

$$E[(d^1x^i - v^i(x,t)h)(d^1x^j - v^j(x,t)h) | \bar{x}(t;h) = \bar{x}] = \sigma^{ij}(t)(h/2) + \pi^{ij}(\bar{x},t)h^2 + o(h^2).$$                    (6.17)

The above result specifies an approximation for the joint probability density $f_h(\bar{x},u,t)$ of the mean position $\bar{x}(t;h)$ and empirical velocity $u(t;h) = d^1x(t;h)/h$ of a reciprocal diffusion. Specifically, we have

$$f_h(x,u,t) \approx \rho(x,t) \frac{\kappa_h^{1/2}(x,t)}{(2\pi)^{n/2}} \exp\left[ -\frac{1}{2}\kappa_{ij,h}(x,t)(u^i - v^i(x,t))(u^j - v^j(x,t)) \right]$$                    (6.18)

with

$$\kappa_h^{ij}(x,t) = \frac{\sigma^{ij}(t)}{2h} + \pi^{ij}(x,t),$$                    (6.19a)

$$\kappa_h(x,t) = \det \kappa_{ij,h}(x,t).$$                    (6.19b)

Thus, the distribution of the velocity $u(t;h)$ given the position $\bar{x}(t;h)$ is approximately Gaussian. Locally in space, it is similar to the Maxwell velocity distribution for a gas in thermal equilibrium. However, one important difference with the Maxwell distribution is that the Maxwell distribution applies to dilute gases, whereas a diffusing particle is subjected to a very large number of collisions with its background bath per unit time, which is reflected by the fact that the leading term of the covariance $\kappa_h^{ij}(x,t)$ is proportional to $1/h$. The correction term $\pi^{ij}(x,t)$, which is the dominant term for dilute gases, is the *stress tensor* of the velocity distribution. In the following, we shall exploit the fact that the density $f_h(x,u,t)$ is completely parametrized by $n(n+1)/2 + n + 1$ functions: the density $\rho(x,t)$, mean velocity vector $v^i(x,t)$ and stress tensor $\pi^{ij}(x,t)$. We have therefore proved the following result.

**Theorem 6.1**: If $x(t)$ is a reciprocal diffusion, for small $h$, the joint probability density $f_h(x,u,t)$ of the mean position $\bar{x}(t;h)$ and velocity $u(t;h)$ has the form (6.18) and (6.19).

In Refs. 7 and 8, the identities (6.14) and (6.17) were introduced in the form of postulates, which in combination with Newton's law, were used to define reciprocal diffusions. Together with Newton's law, these identities can be viewed as an extension to reciprocal diffusions of the postulates

$$E[d^+x^i(t;h) | x(t) = x] = \beta^{i+}(x,t)h + o(h),$$                    (6.20a)

$$E[d^+x^i(t;h)d^+x^j(t;h)\,|\,x(t)=x]=\sigma^{ij}(x,t)h+o(h),\tag{6.20b}$$

where $d^+x(t;h)=x(t+h)-x(t)$ denotes the forward increment of the process $x(t)$, that were employed by Feller[26] to formulate Markov diffusions. One advantage of deriving the velocity distribution, as was done here, instead of introducing it as a postulate, is that we have been able to obtain explicit relations for the mean velocity $v_i(x,t)$ and stress tensor $\pi_{ij}(x,t)$ in terms of the covector potential $A_i(x,t)$ and end-point density $q(x,s;y,t)$. These algebraic relations will be used below in our discussion of the conservation laws of reciprocal diffusions.

## VII. CONSERVATION LAWS

For small $h$, the joint density $f_h(x,u,t)$ of the position and velocity of a reciprocal diffusion is completely parametrized by the density $\rho(x,t)$, mean velocity vector $v^i(x,t)$ and stress tensor $\pi^{ij}(x,t)$. To propagate $f_h(x,u,t)$ in time, we would need to find a closed set of evolution equations for $\rho$, $v^i$ and $\pi^{ij}$. In this section, two conservation laws are derived for the density $\rho$ and momentum density $\rho v^i$, where the second of these equations depends on $\pi^{ij}$. These equations are not sufficient for propagating $f_h(x,u,t)$, since in order to close this system of equations, we would need either a third conservation law or an algebraic relation expressing $\pi^{ij}$ in terms of $\rho$ and $v^i$ only.

To derive the conservation laws, we consider two test functions $f(x,t)$ and $g(x,\xi,t)=g^k(x,t)\xi_k$, where $f$ and the coefficients functions $g^k$ are of class $C^\infty$. Then, if $x_h(t)$ denotes a piecewise linear trajectory passing through the points $x(t_k)$ with $t_k\in D_h$ and satisfying Eq. (3.4) over each trajectory segment, we consider the functions

$$f_h(t)=f(x_h(t),t),\tag{7.1a}$$

$$g_h(t)=g^k(x_h(t),t)\dot{x}_{hk}(t).\tag{7.1b}$$

For a point $t\in D_h$, their centered differences are given by

$$df_h(t)=f_h(t+h/2)-f_h(t-h/2),\tag{7.2a}$$

$$dg_h(t)=g_h(t+h/2)-g_h(t-h/2).\tag{7.2b}$$

Taking into account the approximation (3.5) for the contravariant form of velocity over each trajectory segment, and expanding $x_h(t\pm h/2)$ and $t\pm h/2$ in the vicinity of $\bar{x}$ and $t$, respectively, inside $f_h(t\pm h/2)$ and $g_h(t\pm h/2)$, we obtain

$$df_h(t)=\frac{\partial f}{\partial x^j}(\bar{x},t)d^1x^j+\frac{\partial f}{\partial t}(\bar{x},t)h+\epsilon(t,h),\tag{7.3}$$

$$dg_h(t)=\frac{1}{h}\left[g^k(\bar{x},t)\left(\sigma_{kj}d^2x^j+\frac{\partial\sigma_{kj}}{\partial t}d^1x^j\right)+\frac{\partial g_k}{\partial x^j}(\bar{x},t)\left(d^1x^j\,d^1x^k-\frac{1}{4}d^2x^j\,d^2x^k\right)\right.$$
$$\left.+\frac{\partial g^k}{\partial t}(\bar{x},t)\sigma_{kj}\,d^1x^j\,h\right]+\eta(t,h),\tag{7.4}$$

where the error terms satisfy

$$E[\epsilon(t,h)\,|\,\bar{x}]=O(h^2),\quad E[\eta(t,h)\,|\,\bar{x}]=O(h^2).\tag{7.5}$$

Next, we condition $df_h(t)$ and $dg_h(t)$ with respect to $\bar{x}(t;h)$ and use the fact that the joint conditional density of $d^2x(t;h)$ and $d^1x(t;h)$ given $\bar{x}(t;h)$ takes locally the form of a Gaussian distribution, whose first- and second-order moments can be computed from Newton's law and the velocity distribution obtained in Sec. VI. This gives

$$E[df_h(t)\,|\,\bar{x}=x] = \mathbf{A}f(x,t)h + O(h^2), \tag{7.6}$$

$$E[dg_h(t)\,|\,\bar{x}=x] = \mathbf{B}g(x,t)h + O(h^2), \tag{7.7}$$

where the differential operators $\mathbf{A}$ and $\mathbf{B}$ are defined by

$$\mathbf{A}f(x,t) = \left(v^j\nabla_j f + \frac{\partial f}{\partial t}\right)(x,t), \tag{7.8}$$

$$\mathbf{B}g(x,t) = F_k(x,v(x,t),t)g^k(x,t) + P_{jk}\nabla^j g^k(x,t) + v_k\frac{\partial g^k}{\partial t}(x,t), \tag{7.9}$$

and the *flux of momentum* tensor $P_{jk}(x,t)$ is given by

$$P_{jk}(x,t) \triangleq (\pi_{jk} + v_j v_k)(x,t). \tag{7.10}$$

Since the test functions $f(x,t)$ and $g_k(x,t)$ depend arbitrarily on $x$ and $t$, we can assume

$$f(x,0) = f(x,T) = 0, \tag{7.11a}$$

$$g_k(x,0) = g_k(x,T) = 0, \tag{7.11b}$$

so that

$$0 = \sum_{t\in\mathbf{D}_h} df_h(t) = \sum_{t\in\mathbf{D}_h} E[df_h(t)], \tag{7.12a}$$

$$0 = \sum_{t\in\mathbf{D}_h} dg_h(t) = \sum_{t\in\mathbf{D}_h} E[dg_h(t)]. \tag{7.12b}$$

To evaluate the expected values of $df_h(t)$ and $dg_h(t)$, we can first condition them with respect to $\bar{x}(t;h)$ and then take the expectation of the resulting expression. For example,

$$E[df_h(t)] = E[E[df_h(t)\,|\,\bar{x}]]. \tag{7.13}$$

This amounts to integrating expressions (7.6) and (7.7) against the probability density $p(\bar{x},t;h) = p(\bar{x},t) + O(h)$ of $\bar{x}(t;h)$.

We find

$$0 = \left[\sum_{t\in\mathbf{D}_h} h \int (\mathbf{A}f)p(x,t)dx\right] + O(h), \tag{7.14a}$$

$$0 = \left[\sum_{t\in\mathbf{D}_h} h \int (\mathbf{B}g)p(x,t)dx\right] + O(h). \tag{7.14b}$$

Letting $h\to 0$, the sums over $t$ can be replaced by time integrals, which can be integrated by parts, so that

$$0 = \int\int (\mathbf{A}f)\rho(x,t)dx\,dt = \int\int f(\mathbf{A}^*\rho)(x,t)dx\,dt, \tag{7.15a}$$

$$0 = \int\int (\mathbf{B}g)\rho(x,t)dx\,dt = \int\int g(\mathbf{B}^*\rho)(x,t)dx\,dt, \tag{7.15b}$$

where $\mathbf{A}^*$ and $\mathbf{B}^*$ denote the dual operators of $\mathbf{A}$ and $\mathbf{B}$, respectively. Since the test functions $f(x,t)$ and $g_k(x,t)$ are arbitary, we can conclude that the two conservation laws

$$0 = \mathbf{A}^*\rho = -\left(\frac{\partial\rho}{\partial t} + \nabla^j(\rho v_j)\right), \tag{7.16}$$

$$0 = \mathbf{B}^*\rho = \rho F_k(x,v,t) - \frac{\partial}{\partial t}(\rho v_k) - \nabla^j(\rho P_{jk}) \tag{7.17}$$

hold weakly. The identity (7.16) represents the *conservation of mass*, which is also called the equation of continuity. Similarly, the equation (7.17) represents the *conservation of mean momentum*, since for a small volume element, it equates the forces acting on the element times its mass to the time rate of change of the mean momentum plus the flux of momentum flowing outward across the surface of the element. Note also that, since we are considering a fixed element of volume, the conservation laws are expressed here in Eulerian form.

It is worth observing at this point that the above two conservation laws are actually part of an infinite chain of such laws, which can be obtained as follows. Consider the function $q(x,t;y,t)$ obtained by setting $s=t$ in the end-point density function $q(x,s;y,t)$, i.e., by letting the interval $[s,t]$ shrink to zero. The evolution equations (2.21a) and (2.21b) imply that $q(x,t;y,t)$ satisfies

$$\frac{\partial q}{\partial t}(x,t;y,t) = (\mathbf{H}_{y,t} - \mathbf{H}^*_{x,t})q(x,t;y,t). \tag{7.18}$$

Next, if we define

$$Q(m,z,t) \triangleq q(m-z/2,t;m+z/2,t), \tag{7.19}$$

by performing a Taylor series expansion of $Q(m,z,t)$ in the vicinity of $z=0$, and taking into account the definitions (6.4a) and (6.4b) of $w_i$ and $\pi_{ij}$, we obtain

$$Q(m,z,t) = \rho(m,t)\left[1 + w_i(m,t)z^i + (\pi_{ij} + w_iw_j)(m,t)\frac{z^iz^j}{2} + \tau_{ijk}(m,t)\frac{z^iz^jz^k}{6} + \cdots\right]. \tag{7.20}$$

Performing the change of coordinates $x=m-z/2$, $y=m+z/2$ inside Eq. (7.18), and identifying the successive powers of $b$ on both sides of Eq. (7.18) yields the desired chain of conservation laws. This chain does not close, since the conservation law for each coefficient of Eq. (7.20) contains a divergence term involving the next higher order tensor in the expansion. One exception is the Gaussian case, where because ln $Q(m,z,t)$ is quadratic in $m$ and $z$, the chain closes after the third law, which is given in Ref. 7. However, for non-Gaussian reciprocal diffusions, the conservation laws beyond the first two take a complicated form and are hard to interpret physically.

## VIII. CLOSURE RULES

In this section we consider two closure rules which provide a mechanism for closing the conservation laws of reciprocal diffusions after the first two laws. The two closure rules require that the functions $w_i(x,t)$ and $\pi_{ij}(x,t)$ given by Eqs. (6.4a) and (6.4b) should satisfy

$$w_i(x,t) = \nabla_i S(x,t), \tag{8.1a}$$

$$\pi_{ij}(x,t) = \frac{\epsilon}{4} \nabla_i \nabla_j \ln \rho(x,t), \tag{8.1b}$$

where $S(x,t)$ is an arbitrary function, and $\epsilon = \pm 1$.

As can be seen from Eq. (6.10), the case $\epsilon = 1$ corresponds to the subclass of Markov diffusions. In this case, according to Eq. (6.9), we have

$$S(x,t) = \ln \left( \frac{q_b(x,t)}{q_f(x,t)} \right)^{1/2}, \tag{8.2}$$

and the two conservation laws (7.16) and (7.17) for $\rho$ and $\rho v_k$ are equivalent to the forward and backward evolution equations (2.35a) and (2.35b) for $q_f$ and $q_b$. On the other hand the case $\epsilon = -1$ specifies what we shall call the class of *quantum diffusions*, since in this case, the reciprocal conservation laws are equivalent to Schrödinger's equation.

To prove this, it is convenient to rewrite the conservation laws (7.16) and (7.17) in Lagrangian form as

$$\frac{\partial \rho}{\partial t} + \rho \nabla^j v_j + v_j \nabla^j \rho = 0, \tag{8.3a}$$

$$\rho \frac{\partial v_k}{\partial t} + \rho v_j \nabla^j v_k + \nabla^j (\rho \pi_{jk}) - \rho F_k = 0. \tag{8.3b}$$

Then if we denote

$$R(x,t) \triangleq \ln \rho^{1/2}(x,t), \tag{8.4}$$

by substituting the closure rules (8.1a) and (8.1b) inside Eqs. (8.3a) and (8.3b), we obtain

$$\frac{\partial R}{\partial t} + (\nabla^j S - A^j) \nabla_j R + \frac{\Delta S}{2} - \frac{1}{2} \nabla^j A_j = 0, \tag{8.5}$$

$$\nabla_k I(x,t) = 0, \tag{8.6a}$$

with

$$I(x,t) \triangleq \frac{\partial S}{\partial t} + \frac{1}{2} (\nabla^j S - A^j)(\nabla_j S - A_j) + \phi + \frac{\epsilon}{2} (\nabla^j S \nabla_j S + \Delta R). \tag{8.6b}$$

Provided that $I(x,.) = 0$ at one point in space, the relation (8.6a) implies

$$I(x,t) \equiv 0. \tag{8.7}$$

In the Markov case, i.e., when $\epsilon = 1$, if we introduce the two functions

$$q_f(x,t) = \exp(R-S)(x,t), \quad q_b(x,t) = \exp(R+S)(x,t), \tag{8.8}$$

it is easy to verify that Eqs. (8.5) and (8.7) are equivalent to the forward and backward evolution equations (2.35a) and (2.35b) for $q_f$ and $q_b$.

In the quantum case, i.e., with $\epsilon = -1$, if we introduce the wave function

$$\psi(x,t) = \exp(R(x,t) + iS(x,t)), \tag{8.9}$$

the identities (8.5) and (8.7) correspond, respectively, to the real and imaginary parts of the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(x,t) = \mathbf{H}_Q\psi(x,t), \tag{8.10}$$

where we have set Planck's constant $\hbar = 1$, and

$$\mathbf{H}_Q = \tfrac{1}{2}(-i\nabla^j - A^j(x,t))(-i\nabla_j - A_j(x,t)) + \phi(x,t) \tag{8.11}$$

is the quantum Hamiltonian operator obtained by applying the correspondence principle $p_j \leftrightarrow -i\nabla_j$ to the Hamiltonian $H(x,p,t)$. Note that the wave function $\psi(x,t)$ is consistent with the quantum diffusion $x(t)$ used to construct it, since

$$\rho(x,t) = |\psi(x,t)|^2. \tag{8.12}$$

At this stage, the correspondence between quantum diffusions and quantum mechanical processes is purely formal, since we have not yet shown how to select an end-point density $q(x_0,x_T)$ such that the two closure rules (8.1a) and (8.1b) are satisfied with $\epsilon = -1$. This topic will be examined in Ref. 14, where a complete formulation of stochastic mechanics based on quantum diffusions is presented.

Finally, observe that the closure rule (8.1a) is just a requirement that $w_i(x,t)$ should be a gradient function. But it is easy to verify that the momentum conservation law implies that if $w_i$ is a gradient at $t=0$, it remains a gradient for all times. Thus, the only truly constraining closure rule is (8.1b).

## IX. CONCLUSIONS

In this paper we have obtained a complete characterization of the dynamics and kinematics of reciprocal diffusions. This was accomplished by employing a "stochastic quantization" procedure, where the Hamiltonian of a dynamical system was used to specify a generalized heat operator, which in turn yielded the finite joint densities of a family of reciprocal processes, where each member in the family differs from another one only by the end-point probability density $q(x_0,x_T)$ assigned to it. A path integral representation of the heat operator Green's function was obtained, which was then used to derive a stochastic Newton law and a distribution for the velocity of diffusing particles. These results were employed to derive conservation of mass and momentum laws which are closed for the subclasses of Markov and quantum diffusions.

The class of quantum diffusions is studied in further detail in Ref. 14, where it is used to develop a new version of stochastic mechanics which is free of several of the inconsistencies with quantum mechanics affecting Nelson's theory, which was based on Markov diffusions.

The results presented here are still incomplete in several respects. One issue that needs further examination concerns the extension of the above results to curved space, i.e., to the case when the metric $\sigma_{ij}$ depends on both space and time. This will require finding a more general

stochastic quantization rule, and obtaining a path integral representation of $G$ in curved space. Both issues are fraught with pitfalls, but they have been worked out in quantum mechanics.

Another element missing from the above study of reciprocal diffusions is a description in terms of stochastic differential equations. Since the stochastic Newton law is expressed in terms of second-order differences of the diffusion process $x(t)$, it seems natural to interpret it as a second-order stochastic differential equation, as was done for the Gaussian case in Ref. 25. However, one feature that makes the development of such a theory difficult is that, as shown in Ref. 25 in the Gaussian case, the driving noise for the second-order differential equation is locally correlated, with a correlation depending on the forces acting on the system.

## ACKNOWLEDGMENT

[1] E. Schrödinger, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl. **1931**, 144; Ann. Inst. Henri Poincaré **2**, 269 (1932).

[2] S. Bernstein, Sur les liaisons entre les grandeurs aléatoires, in Proc. Int. Cong. of Math., Zurich, Switzerland (1932).

[3] P. Lévy, "A special problem of Brownian motion and a general theory of Gaussian random functions," in *Proceedings of the 3rd Berkeley Symposium on Probability and Statistics*, Vol. 2, edited by J. Neyman (University of California, Berkeley, 1956).

[4] B. Jamison, Z. Wahrscheinlichkeitstheorie verw. Gebiete **30**, 65 (1974); **32**, 323 (1975).

[5] E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton University, Princeton, NJ, 1967) and *Quantum Fluctuations* (Princeton University, Princeton, NJ, 1985).

[6] P. Blanchard, P. Combe, and W. Zheng, *Mathematical and Physical Aspects of Stochastic Mechanics*, Lecture Notes in Physics, Vol. 281 (Springer-Verlag, Berlin, 1987).

[7] A. J. Krener, Stochastics **24**, 393 (1988).

[8] A. J. Krener (unpublished manuscript, 1988).

[9] J. M. C. Clark, "A local characterization of reciprocal diffusions," in *Applied Stochastic Analysis*, Vol. 5, edited by M. H. A. Davis and R. J. Elliott (Gordon and Breach, New York, 1991).

[10] B. Øksendal, *Stochastic Differential Equations*, 2nd. ed. (Springer-Verlag, Berlin, 1989).

[11] J. C. Zambrini, Phys. Rev. A **33**, 1532 (1986); J. Math. Phys. **27**, 2307 (1986); Phys. Rev. A **35**, 3631 (1987).

[12] K. Kime and A. Blaquière, "From two stochastic optimal control problems to the Schrödinger equation," in *Modeling and Control of Systems in Engineering, Quantum Mechanics, Economics and Biosciences*, edited by A. Blaquière, Lecture Notes in Control and Information Sciences, Vol. 121, (Springer-Verlag, Berlin, 1989); A. Walkobinger, J. Math. Phys. **30**, 2943 (1989).

[13] M. Kac, "On some connections between probability theory and differential and integral equations," in *Proceedings of the 2nd Berkeley Symp. on Math. Stat. and Prob.*, edited by J. Neyman (University of California, Berkeley, 1951); *Probability Theory and Related Topics in Physical Sciences* (Interscience, London, 1959).

[14] B. C. Levy and A. J. Krener (in preparation).

[15] A. Friedman, *Partial Differential Equations of Parabolic Type* (Prentice Hall, Englewood Cliffs, NJ, 1964).

[16] R. Fortet, J. Math. Pures et Appl. **9**, 83 (1940); A. Beurling, Ann. Math. **72**, 189 (1960).

[17] M. Nagasawa, Prob. Theory and Related Fields **82**, 109 (1989); "Can the Schrödinger equation be a Boltzmann equation?," in *Diffusion Processes and Related Problems in Analysis*, Vol. 1, edited by M. A. Pinsky (Birkhäuser, Boston, 1990).

[18] K. Itô and H. P. McKean, *Diffusion Processes and their Sample Paths* (Springer-Verlag, Berlin, 1965).

[19] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965); L. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981); M. S. Marinov, Phys. Rep. **60**, 1 (1980).

[20] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).

[21] D. Durr and A. Bach, Commun. Math. Phys. **60**, 153 (1978); Y. Takahashi and S. Watanabe, *The probability functionals (Onsager–Machlup functions) of diffusion processes*, in Lecture Notes in Mathematics, Vol. 851 (Springer-Verlag, Berlin, 1981).

[22] R. Graham, Z. Phys. B **26**, 281 (1977).

[23] B. S. DeWitt, Rev. Mod. Phys. **29**, 377 (1957); D. W. McLaughlin and L. Schulman, J. Math. Phys. **12**, 2520 (1971).

[24] S. A. Molchanov, Russ. Math. Surveys **30**, 1 (1975); R. Azencott, "Densité de diffusions en temps petit: dévelop-

pements asymptotiques," in *Seminaire de Probabilités XVIII 1982/1983*, Lecture Notes in Math., Vol. 1059 (Springer-Verlag, Berlin, 1984).

[25] B. C. Levy, R. Frezza, and A. J. Krener, IEEE Trans. Automat. Cont. **35**, 1013 (1990); A. J. Krener, R. Frezza, and B. C. Levy, Stochastics and Stochastic Rep. **34**, 29 (1991).

[26] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2 (Wiley, New York, 1966).