

Conservative Diffusions

Eric A. Carlen

Department of Physics, Princeton University, Princeton, NJ 08544, USA

Abstract. In Nelson's stochastic mechanics, quantum phenomena are described in terms of diffusions instead of wave functions. These diffusions are formally given by stochastic differential equations with extremely singular coefficients. Using PDE methods, we prove the existence of solutions. This result provides a rigorous basis for stochastic mechanics.

Introduction

In Nelson's stochastic mechanics quantum phenomena are described in terms of diffusions instead of wavefunctions. The class of diffusions considered here arises in stochastic mechanics; and in order to motivate our theorems, we provide a brief description of the theory. There is unfortunately a certain amount of technical material involved, so our description here must be illustrative rather than precise. In the main body of the paper we will be precise.

Consider a point particle in \mathbb{R}^3 moving under the influence of a potential $V(x)$. When we specify what we mean by moving—the kinematics, and what we mean by influence of a potential—the dynamics, we will have a mechanics.

The kinematical proposal is that we describe the motion of our particle with a Markovian diffusion $\xi(t)$ in \mathbb{R}^n . (We define all terms from diffusion theory we use at the start of Sect. I.) To get agreement with quantum mechanics, we must specify the size of the fluctuations; we require that on any interval $[a, b]$, the quadratic variation of our process is $(\hbar/2m)(b - a)$ a.s., where \hbar is Planck's constant over 2π and m is the mass of the particle. Notice this has the correct units for a quadratic variation, and henceforth set $\hbar = m = 1$. We will not discuss the possible physical natures of these fluctuations—see Nelson [1] for a conjecture—but will simply remark that if we assume them to be manifestations of an isotropic translation invariant phenomenon, then the second order part of the generator of our diffusion must be $\frac{1}{2}\Delta$. So, the diffusions to be considered are those with a time dependent generator of the form $\frac{1}{2}\Delta + b(x, t) \cdot \nabla$.

The dynamics is given by the Guerra–Morato–Nelson variational principle. Fix a finite time interval $[0, T]$, an initial density $\rho_0(x)$ and an initial forward drift $b_0(x)$. Consider the class of diffusions with generator of the form $\frac{1}{2}\Delta + b(x, t) \cdot \nabla$, where $b(x, 0) = b_0(x)$ and with $\xi(0)$ having density $\rho_0(x)$. For such a diffusion $\xi(t)$,

define its action on $[0, T]$ to be:

$$\mathcal{A}(\xi) \equiv E \left(\int_0^T \left(\frac{1}{2} \dot{\xi}(t)^2 - V(\xi(t)) \right) dt \right) \equiv E \left(\int_0^T L(t) dt \right). \tag{1}$$

Although the sample paths of $\xi(t)$ are nowhere differentiable, a theorem of Nelson [1] shows that the above action has an unambiguous meaning in a variational context. Roughly speaking, $\mathcal{A}(\xi)$ is the sum of a nice function of ξ and an infinite constant independent of ξ . The infinite constant drops out on taking variations. Then a theorem of Guerra and Morato [2] shows that a diffusion $\xi(t)$ is a stationary point of the above action precisely when there is a solution ψ of the Schroedinger equation

$$i \frac{\partial \psi}{\partial t} = \left(-\frac{1}{2} \Delta + V(x) \right) \psi, \tag{2}$$

so that the forward drift $b(x, t)$ of $\xi(t)$ is given by

$$b(x, t) = \text{Im} \frac{\nabla \psi(x, t)}{\psi(x, t)} + \text{Re} \frac{\nabla \psi(x, t)}{\psi(x, t)}, \tag{3}$$

and so that the density $\rho(x, t)$ of $\xi(t)$ is given by

$$\rho(x, t) = |\psi(x, t)|^2. \tag{4}$$

Note that the same potential $V(x)$ appears in both (1) and (2). We have now specified both the kinematics and dynamics of stochastic mechanics.

Equation (4) says that stochastic mechanics and ordinary quantum mechanics make the same predictions for the probability of finding our particle in any given place at any given time. The simple example discussed here by no means exhausts the scope of stochastic mechanics. For systems with quite general configuration spaces and potentials, stochastic mechanics still makes sense and still makes the same predictions as ordinary quantum mechanics for the probability of finding the system in any given configuration at any given time. Since all measurement processes ultimately consist of configuration measurements, the two theories are experimentally indistinguishable.

To conclude, a rather direct transcription of Lagrangian mechanics into stochastic terms leads one to the Schroedinger equation; and in fact to a description of quantum phenomena experimentally indistinguishable from that given by the Schroedinger equation, but in terms of actual point particle motions. The formal equations for the diffusions of stochastic mechanics were first discovered by Fényes [3]. They were rediscovered by Nelson [4], who first explained their dynamical significance using a stochastic version of Newton's equation. The variational approach discussed above was given by Guerra and Morato in [2]; other variational approaches have been given by Yasue [5] and Zambrini [6]. For a comprehensive discussion, see the book [1] by Nelson.

In this paper we treat the following question: given a potential $V(x)$ and a solution $\psi(x, t)$ of the Schroedinger equation for this potential, does there exist a diffusion with initial density $|\psi(x, t)|^2$ and generator $\frac{1}{2} \Delta + b(x, t) \cdot \nabla$, where $b(x, t)$

is given by (3)? Our result is: yes, if $V(x)$ is a Rellich class potential (this is almost no restriction at all; see Sect. II for its meaning) and if $\psi(x, 0)$ has finite kinetic energy (i.e. $\|\nabla\psi(x, 0)\|^2 < \infty$.)

This is not at all a trivial problem. If such a diffusion exists, it must exhibit all the conservative behavior that the corresponding solution of the Schroedinger equation exhibits. The processes we construct here, while still continuous path Markov processes, are qualitatively different from the dissipative diffusions—such as Brownian motion—that one usually considers. Following Nelson, we call them conservative diffusions.

The problem of constructing a diffusion with a given initial density and a given generator has been treated by many authors. However, all known methods are designed to treat the case where $b(x, t) \cdot \nabla$ is close to being a small perturbation of $\frac{1}{2}\Delta$. This is reflected in the fact that these methods produce measures on path space which are simply absolutely continuous transformations of the Wiener measure with the given initial density. In our case $\nabla\psi$ is only known to be square integrable, and ψ may have zeros, so the drift vector field $b(x, t)$ given by (3) can be a very singular object indeed. It cannot be split off the Laplacian and treated as a small perturbation in any way. See Meyer and Zheng [7] for more discussion of this problem. There are some previous results in special cases. Albeverio and Hoegh-Krohn [8] and Carmona [9] have treated the case of stationary states, and Nelson [10] has treated the case of a compact configuration space under some smoothness assumptions.

We now summarize the paper. Section I presents diffusion theory from a time symmetric point of view and then outlines the strategy of our construction. Section II contains the proof of some regularity properties of finite energy solutions of the Schroedinger equation with a Rellich class potential. These will be needed in the following sections. Section III contains the PDE result on which our approach is based. Section IV contains the actual construction of our conservative diffusions.

I. Diffusions and Their Martingale Equations

We begin with some terminology and notation. Let Ω_c be the space of continuous functions $\omega: \mathbb{R} \rightarrow \mathbb{R}^n$ given the topology of uniform convergence on compacts. This makes it a complete separable metrizable space, and we will call it *trajectory space*. (We reserve the term path space for a different space occurring later.) There is a distinguished class of functions on Ω_c : the *t-configuration function* $\xi(t)$ is defined by $\xi(t): \omega \mapsto \omega(t) \equiv \xi(t, \omega)$. It is not hard to show that the Borel field \mathcal{B} on Ω_c is given by $\mathcal{B} = \sigma\{\xi(t): t \in \mathbb{R}\}$; see [11]. Certain sub σ algebras of \mathcal{B} are of special interest. The *past at time t*, \mathcal{P}_t , is given by $\mathcal{P}_t = \sigma\{\xi(s): s \leq t\}$. The intuitive meaning of this is that a Borel function on Ω_c is \mathcal{P}_t measurable precisely when it only depends on configurations up till time t . Similarly the *future at time t*, \mathcal{F}_t , is given by $\mathcal{F}_t = \sigma\{\xi(u): u \geq t\}$, and the *present at time t*, \mathcal{N}_t , is given by $\mathcal{N}_t = \sigma\{\xi(t)\}$. (The \mathcal{N} is for now.)

Suppose we are given a \mathcal{B} -probability measure Pr on Ω_c . Then Borel functions f become random variables, and we denote their expectations—provided they are integrable—by Ef . If \mathcal{A} is a sub σ algebra of \mathcal{B} , and if $f \in L^1(Pr)$, then we will

denote the conditional expectation of f given \mathcal{A} by $E\{f|\mathcal{A}\}$ except when $\mathcal{A} = \mathcal{N}_t$ for some t ; then we will just write $E_t f$.

When trajectory space is made into a probability space in this way, $t \mapsto \xi(t)$ becomes a stochastic process which we will call the *configuration process*. Recall that to say $\xi(t)$ is Markovian means that whenever $f \in \mathcal{P}_t$ and $g \in \mathcal{F}_t$ with f, g , and fg integrable, $E_t(fg) = E_t(f) \cdot E_t(g)$.

We will say that under Pr the configuration process is a *smooth diffusion* with coefficients $\sigma^{ij}(x, t)$, $b^i(x, t)$, and $b_*^i(x, t)$, $1 \leq i, j \leq n$ in case

- (i) The configuration process is Markovian.
- (ii) The σ^{ij} are all smooth bounded functions where the $\sigma^{ij}(x, t)$ are the entries of a positive definite matrix. The b^i and the b_*^i are all smooth functions bounded by $c(1 + |x|)$ for some constant c .
- (iii) The following limits exist for any $f \in C_0^\infty(\mathbb{R}^n)$:

$$\lim_{h \rightarrow 0^+} \frac{1}{h} E_t(f(\xi(t+h)) - f(\xi(t))) = \sum_{i,j} \left(\frac{1}{2} \sigma^{ij} \frac{\partial^2}{\partial x^i \partial x^j} + b^i \frac{\partial}{\partial x^i} \right) f(\xi(t)), \tag{1}$$

$$\lim_{h \rightarrow 0^+} \frac{1}{h} E_t(f(\xi(t)) - f(\xi(t-h))) = \sum_{i,j} \left(-\frac{1}{2} \sigma^{ij} \frac{\partial^2}{\partial x^i \partial x^j} + b_*^i \frac{\partial}{\partial x^i} \right) f(\xi(t)). \tag{2}$$

Notice, as emphasized by Nelson, that the definition of a diffusion is time symmetric. It may seem that we have contrived this to be true by adding (2) to the definition. However, it is not hard to see that if (1) holds and $\xi(0)$ has a smooth density free of zeros, then there is a vector field with components $b_*^i(x, t)$, compatible with (ii), so that (2) is satisfied. We have just rephrased the usual definition—in a slightly redundant way—to bring out its symmetry. An easy example is the Wiener process with initial density $\rho_0(x) = (2\pi)^{-n/2} \exp(-x^2/2)$. Then on $[0, T]$ we have $\sigma^{ij}(x, t) = \delta^{ij}$, $b^i(x, t) = 0$, $b_*^i(x, t) = -x^i/(1+t)$.

The matrix with entries $\sigma^{ij}(x, t)$ is called the *diffusion matrix*. It determines the size of the fluctuations in the sense that $\lim_{n \rightarrow \infty} \sum_{m=1}^{2^n} (\xi^i(t_m) - \xi^i(t_{m-1})) \cdot (\xi^j(t_m) - \xi^j(t_{m-1})) = \int_a^b \sigma^{ij}(\xi(s)) ds$ a.s., where the sum is over the endpoints of the n^{th} dyadic partition of $[a, b]$. This quantity is called the *quadratic variation* of $\xi(t)$ on $[a, b]$. As indicated in the introduction, we will be concerned with the case where σ^{ij} identically equals δ^{ij} , and we henceforth restrict our attention to it.

This done, the second order operator in (1) becomes $\frac{1}{2} \Delta + b(x, t) \cdot \nabla$ and is called the (*forward*) *generator* of $\xi(t)$. The vector field $b(x, t)$ is called the *forward drift*; it gives the best possible prediction based on information in the past at time t of how the configuration will change just after time t . Similarly $b_*(x, t)$ is called the *backward drift*, and it has the analogous interpretation.

Suppose $\xi(t)$ has a smooth density with respect to Lebesgue measure on \mathbb{R}^n —by this we mean that the image of Pr under $\xi(t)$ has such a density. Then it is known that this property holds for all times in the interval considered if $\xi(t)$ is a smooth diffusion. (Use the regularity of b to work forward, and use the regularity of b_* to work backward.) Assume this is the case. Then we have the following relation

between b, b_* , and ρ , the smooth density:

$$b_*(x, t) = b(x, t) - \nabla \log(\rho(x, t)). \tag{3}$$

We will find the odd and even combinations of b and b_* useful. Following Nelson we define the *osmotic velocity* $u(x, t)$ by

$$u(x, t) = \frac{1}{2}(b(x, t) - b_*(x, t)), \tag{4}$$

and we define the *current velocity* $v(x, t)$ by

$$v(x, t) = \frac{1}{2}(b(x, t) + b_*(x, t)). \tag{5}$$

It is then a theorem that u, v , and ρ are related by the following equations:

$$u(x, t) = \frac{1}{2} \nabla \log(\rho(x, t)), \tag{6}$$

$$\frac{\partial \rho}{\partial t}(x, t) = -\nabla \cdot (v(x, t)\rho(x, t)). \tag{7}$$

For the results quoted above and for a more detailed discussion of diffusion from a time symmetric point of view, see [1]. Note that (6) follows from (3), (4) and (5). Equation (7) is the continuity equation for ρ , and it is fundamental in what follows.

We now turn to the problem of constructing a diffusion with given coefficients. We will use a somewhat nonstandard PDE method to solve this problem in our case of singular coefficients. To clarify our aim, we briefly discuss our approach in the simple smooth case.

In the following, assume Pr is a measure on trajectory space making the configuration process a smooth diffusion with backward drift $b_*(x, t)$. We will work on a fixed finite time interval $[0, T]$.

We say that a stochastic process $t \rightarrow F_t \in L^1(Pr)$ is a *backward martingale* in case each F_t is \mathcal{F}_t measurable and $E\{F_s | \mathcal{F}_t\} = F_t$ for all $0 \leq s \leq t \leq T$. The more familiar (forward) martingales have the analogous definition.

Let $f(x, t)$ be a smooth bounded function on $\mathbb{R}^n \times [0, T]$. Then the process $f(\xi(t), t)$ is a backward martingale precisely when $f(x, t)$ satisfies the *backward martingale equation*:

$$\frac{\partial f}{\partial t}(x, t) = (\frac{1}{2}\Delta - b_*(x, t) \cdot \nabla)f(x, t). \tag{8}$$

This follows essentially from (2); see [1] for the proof.

Equation (8) is parabolic; and for regular b_* , standard theorems guarantee the existence, uniqueness and regularity of Markovian transition function which generates its bounded solutions. If we denote this transition function by $p_*(y, t; x, s)$, and if $f(x, t)$ satisfies (8), then

$$f(y, t) = \int p_*(y, t; x, s)f(x, s)dx, \quad 0 \leq s \leq t \leq T. \tag{9}$$

But since in this case $f(\xi(t), t)$ is a backward martingale, we have by definition:

$$f(\xi(t), t) = E_t f(\xi(s), s), \quad 0 \leq s \leq t \leq T. \tag{10}$$

Now if $\check{p}_*(y, t; x, s)$ is the backward transition function of $\xi(t)$ —see [1]—we may

compute future conditional expectations in terms of it as follows:

$$f(\xi(t), t) = \int \tilde{p}_*(\xi(t), t; x, s) f(x, s) dx. \tag{11}$$

Comparing the last three equations, we see that the transition function for the backward martingale equation is precisely the backward transition function for our diffusion. Therefore, if we can solve (8) for its transition function, we can use this and the given final density $\rho(x, T)$ to construct a measure Pr on Ω_c by a well known method. It is then easy to see that Pr is the unique measure making $\xi(t)$ a smooth diffusion with backward generator $(\frac{1}{2}\Delta + b_* \cdot \nabla)$ and final density $\rho(x, T)$. In particular, this construction recovers Pr .

The above discussion and our ultimate goal motivate the following problem: Let $\psi(x, t)$ be a solution of the Schroedinger equation for some potential $V(x)$. Define:

$$u(x, t) = \begin{cases} \operatorname{Re} \frac{\nabla \psi(x, t)}{\psi(x, t)} & \psi(x, t) \neq 0 \\ 0 & \psi(x, t) = 0, \end{cases} \tag{12}$$

$$v(x, t) = \begin{cases} \operatorname{Im} \frac{\nabla \psi(x, t)}{\psi(x, t)} & \psi(x, t) \neq 0 \\ 0 & \psi(x, t) = 0. \end{cases} \tag{13}$$

Note that $u(x, t) = \frac{1}{2} \nabla \log |\psi(x, t)|^2$, where the latter is defined. The process we want to construct is supposed to have backward drift b_* given by $b_* = v - u$, where u and v are given by (12) and (13). The corresponding backward martingale equation is then

$$\frac{\partial f}{\partial t}(x, t) = (\frac{1}{2}\Delta - (v - u) \cdot \nabla) f(x, t). \tag{14}$$

The problem is to solve (14) for its transition function. (Actually, we only need to find the Markovian operators $P_{t,s}$ given by $P_{t,s} f(x) = \int p_*(x, t; w, s) f(w, s) dw$, where p_* is the transition function.)

To solve this problem, we impose the following minimal regularity conditions on u, v , and ρ :

The Finite Action Condition

This is the condition that

$$\int_0^T (u^2 + v^2) \rho(x, t) dx dt < \infty, \tag{15}$$

where by definition

$$\rho(x, t) = |\psi(x, t)|^2. \tag{16}$$

A simple computation using the formulas (12), (13) and (16) shows the quantity in

(15) is given by $\int_0^T \|\nabla\psi(x, t)\|^2 dt$. That is, we are requiring the time integral of the quantum mechanical kinetic energy—which is an action integral—to be finite.

The Weak Continuity Equation

Whenever f is a bounded function on \mathbb{R}^n with bounded continuous first derivatives, we require that $t \mapsto \int f(x)\rho(x, t)dx$ be differentiable a.e. and that

$$\frac{d}{dt} \int f(x)\rho(x, t)dx = \int (v(x, t) \cdot \nabla f(x))\rho(x, t)dx \quad \text{a.e. } dt \tag{17}$$

This is called the weak continuity equation because it's what one gets using f as a test function and integrating by parts on the continuity equation (7).

The main result of Sect. III is that we can solve (8) for its transition function if these two conditions hold. First however in Section II, we show that if ψ is a finite energy solution of the Schroedinger equation for a Rellich class potential, then u, v and ρ defined in terms of ψ do in fact satisfy these conditions.

We close this section by comparing our PDE approach to constructing diffusions with the usual one. With $f(x, t)$ as before, $f(\xi(t), t)$ is a forward martingale precisely when $f(x, t)$ satisfies the forward martingale equation $\frac{\partial f}{\partial t}(x, t) = (-\frac{1}{2}\Delta - b \cdot \nabla)f(x, t)$, which is usually called the backward diffusion equation. (There is no such relation between the backward martingale equation and the forward diffusion equation.) One can produce the forward transition function of the process we seek by solving the forward martingale equation; indeed, this is the usual approach. However, the forward martingale equation is antiparabolic as it stands, and so one must work backwards from a given final time condition to solve it. This would be inconvenient for our purposes in Section III.

II. Finite Energy Solutions of the Schroedinger Equation

A solution $\psi(x, t)$ of the Schroedinger equation

$$i \frac{\partial}{\partial t} \psi(x, t) = (-\frac{1}{2}\Delta + V(x))\psi(x, t) \text{ with } \psi(x, 0) = \psi_0(x) \tag{1}$$

fills entirely different roles in ordinary quantum mechanics and in stochastic mechanics. In the first case ψ gives the complete specification of the state of the system. In the second case ψ is of interest merely because it tells us what the density and drifts of the corresponding process are, as in Eqs. (12), (13) and (16) of Sect. II.

Somewhat ironically, in the second more limited physical role, we must ask more regularity of ψ . In particular, we will need to know that the finite action condition holds, and that the weak continuity equation is satisfied. These conditions are irrelevant in ordinary quantum mechanics. To obtain this regularity, we must impose conditions on $\psi_0(x)$ and $V(x)$ in (1). Our main result in this section—Theorem 2.1—is that ψ possesses all necessary regularity properties if

$\|\nabla\psi_0\|^2 < \infty$, and if $V(x)$ is a Rellich class potential. Further useful regularity holds if $\|x\psi_0\|^2 < \infty$.

We will denote $-\frac{1}{2}\Delta$ by H_0 , and we will denote its domain (as a selfadjoint operator on $L^2(\mathbb{R}^n)$) by $\mathcal{D}(H_0)$. Its form domain, $\mathcal{Q}(H_0)$, is then just the space of $\psi \in L^2(\mathbb{R}^n)$ such that $\|\nabla\psi\|^2 < \infty$. We can express our condition on ψ_0 as $\psi_0 \in \mathcal{Q}(H_0)$.

A potential $V(x)$ is called a *Rellich class* potential if, as a multiplication operator on $L^2(\mathbb{R}^n)$ the domain of V contains $\mathcal{D}(H_0)$, and for some a and b in \mathbb{R} with a strictly less than 1, the following inequality holds for all φ in $\mathcal{D}(H_0)$:

$$\|V\varphi\| \leq a\|H_0\varphi\| + b\|\varphi\|.$$

This is a weak restriction on the potential. For instance, Coulomb potentials are Rellich class. In \mathbb{R}^3 , Rellich class consists precisely of $L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$. We will always take V to be real so that it is symmetric. The Kato–Rellich theorem asserts that when V is Rellich class, $H \equiv H_0 + V$ is a semibounded selfadjoint operator with domain $\mathcal{D}(H_0)$. We can then use Stone’s theorem to write the solution to (1) as $\psi(x, t) = \exp(-itH)\psi_0(x)$. Proofs for these assertions are contained in [12].

Next we quote a theorem of Bochner and Von Neumann which we will use here and in Sect. IV.

Bochner-Von Neumann Measurability Theorem. *Let (X, dm) be a separable measure space, and let $t \mapsto f_t$ be a map from $[0, T]$ into $L^2(dm)$. If $t \mapsto f_t$ is weakly measurable and $t \mapsto \int |f_t|^2 dm$ is in $L^1(0, T)$, then there exists a jointly measurable function f on $X \times [0, T]$ such that for almost every t , $\int |f(x, t) - f_t(x)|^2 dm = 0$.*

This function is clearly unique up to equivalence a.e. $dm \times dt$, and we call it the jointly measurable version of f_t . For the proof see [13] and [14].

We are now ready to give Theorem 2.1.

Theorem 2.1 *Let V be a Rellich class potential and let $\psi_0 \in \mathcal{Q}(H_0)$. Let $\psi_t = \exp(-itH)\psi_0$. Then:*

- (i) *For all t , $\psi_t \in \mathcal{Q}(H_0)$, and $t \mapsto \|\nabla\psi_t\|^2$ is continuous.*
- (ii) *There are unique jointly measurable functions $\psi(x, t)$ and $\nabla\psi(x, t)$ such that $\psi(\cdot, t) = \psi_t$ and $\nabla\psi(\cdot, t) = \nabla\psi_t$ a.e. dx .*
- (iii) *With u, v and ρ defined in terms of $\psi(x, t)$ and $\nabla\psi(x, t)$ as in (12), (13) and (16) of Sect. II, we have that for each finite interval $[0, T]$, there exists a constant $M < \infty$ depending only on T and such that*

$$\int (u^2(x, t) + v^2(x, t))\rho(x, t)dx < M \text{ for a.e. } t \in [0, T]. \tag{2}$$

Also, the weak continuity equation holds: for all real bounded functions f on \mathbb{R}^n with bounded continuous first derivatives, $t \mapsto \int f(x)\rho(x, t)dx$ is continuously differentiable, and

$$\frac{d}{dt} \int f(x)\rho(x, t)dx = \int (v(x, t) \cdot \nabla f(x))\rho(x, t)dx. \tag{3}$$

- (iv) *The density $\rho(x, t)$ has finite second moments, uniformly bounded in any compact interval, provided only that $\int x^2\rho(x, 0)dx$ is finite.*

Proof. We recall that the Kato–Rellich theorem asserts that H is selfadjoint on $\mathcal{D}(H_0)$ and bounded below by $-b/(1-a) \equiv -(\lambda-1)$. Then $\mathcal{D}(H_0)$ is a Banach space under each of the norms $\|\cdot\|_0$ and $\|\cdot\|_1$, where $\|\varphi\|_0 = \|(H_0 + \lambda)\varphi\|$ and $\|\varphi\|_1 = \|(H + \lambda)\varphi\|$. By the closed graph theorem, these are equivalent norms, and so there exist constants c_1 and c_2 (both positive) with which:

$$c_1 \|\varphi\|_1 \leq \|\varphi\|_0 \leq c_2 \|\varphi\|_1 \text{ for all } \varphi \in \mathcal{D}(H_0). \quad (4)$$

The Heinz theorem [15] then asserts that for any $0 < \alpha < 1$, $\mathcal{D}((H_0 + \lambda)^\alpha) = \mathcal{D}((H + \lambda)^\alpha) \equiv \mathcal{D}^\alpha$, and that

$$c_1^\alpha \|(H + \lambda)^\alpha \varphi\| \leq \|(H_0 + \lambda)^\alpha \varphi\| \leq c_2^\alpha \|(H + \lambda)^\alpha \varphi\| \text{ for all } \varphi \in \mathcal{D}^\alpha. \quad (5)$$

We also observe that for any two $a, b > 0$,

$$(1/2^{1/2})(a^{1/2} + b^{1/2}) \leq (a + b)^{1/2} \leq (a^{1/2} + b^{1/2}).$$

Applying this inequality and the spectral theorem, we deduce:

$$(1/2^{1/2})\|(H_0^{1/2} + \lambda^{1/2})\varphi\| \leq \|(H_0 + \lambda)^{1/2}\varphi\| \leq \|(H_0^{1/2} + \lambda^{1/2})\varphi\|. \quad (6)$$

Together, (5) and (6) imply that φ is in $\mathcal{D}(H_0^{1/2})$ precisely when it is in $\mathcal{D}^{1/2}$. Of course, $\mathcal{D}(H_0^{1/2}) = \mathcal{D}(H_0)$. Since $(H + \lambda)^{1/2}$ commutes with $\exp(-itH)$, it's easy to see $\mathcal{D}^{1/2}$ is invariant under $\exp(-itH)$, and hence so is $\mathcal{D}^{1/2}(H_0)$. Also

$$\begin{aligned} & \|(H + \lambda)^{1/2}(\exp(-isH)\varphi - \exp(-itH)\varphi)\| \\ &= \|(\exp(-isH) - \exp(-itH))(H + \lambda)^{1/2}\varphi\|, \end{aligned}$$

and so $t \mapsto \psi_t$ is continuous into $\mathcal{D}^{1/2}$ for ψ_0 in $\mathcal{D}^{1/2}$. Inequalities (5) and (6) again imply that then $t \mapsto \psi_t$ is continuous into $\mathcal{D}(H_0)$ for ψ_0 in $\mathcal{D}(H_0)$. Since $\psi_t \rightarrow \|\nabla\psi_t\|^2$ is clearly continuous on $\mathcal{D}(H_0)$, we have proved (i).

Given (i), the Bochner–Von Neumann theorem immediately implies (ii).

The first part of (iii) is easy; we have already observed that for almost every t ,

$$\int (u^2(x, t) + v^2(x, t))\rho(x, t)dx = \|\nabla\psi_t\|^2.$$

By (i), the right-hand side is continuous in t and hence bounded on any compact interval.

Now, $H: \mathcal{D}^{1/2} \rightarrow \mathcal{D}^{-1/2}$ is continuous, and since $\psi_0 \in \mathcal{D}^{1/2}$, $i(d/dt)\psi_t = H\psi_t$, where the derivative exists strongly in $\mathcal{D}^{-1/2}$. (This is the completion of $L^2(\mathbb{R}^n)$ with the inner product $(f, g)_{-1/2} \equiv \langle (H + \lambda)^{-1/2}f, (H + \lambda)^{-1/2}g \rangle$.) If f is any real function on \mathbb{R}^n which is bounded and has continuous bounded first derivatives, multiplication by f is a bounded operation on $\mathcal{D}(H_0^{1/2}) = \mathcal{D}^{1/2}$. We let f denote this multiplication operator as usual. It follows that $t \mapsto \langle f\psi_t, \psi_t \rangle$ is continuously differentiable and $(d/dt)\langle f\psi_t, \psi_t \rangle = \langle -iH\psi_t, f\psi_t \rangle + \langle f\psi_t, iH\psi_t \rangle = \frac{1}{2}\langle i\Delta\psi_t, f\psi_t \rangle + \frac{1}{2}\langle f\psi_t, i\Delta\psi_t \rangle = (i/2)(\langle \nabla\psi_t, (\nabla f)\psi_t \rangle - \langle (\nabla f)\psi_t, \nabla\psi_t \rangle)$. By (ii) this is, for almost every t , equal to:

$$\begin{aligned} & \int (1/2i) \left(\frac{\nabla\psi(x, t)}{\psi(x, t)} - \frac{\nabla\psi^*(x, t)}{\psi^*(x, t)} \right) \cdot \nabla f(x)\rho(x, t)dx \\ &= \int v(x, t) \cdot \nabla f(x)\rho(x, t)dx. \end{aligned}$$

This proves (iii), and it remains to prove the assertion about the second moments of ρ .

We will use a monotone convergence argument. Let $g_1(x)$ be a real valued function on \mathbb{R} which is smooth, monotone increasing, and satisfies: $g_1(x) = -2$ for $x \leq -2$, $g_1(x) = x$ for $-1 \leq x \leq 1$ and $g_1(x) = 2$ for $x \geq 2$. It's easy to see we can choose g_1 so that—using primes to denote derivatives—we have $xg_1''(x) \leq 0$; we do so. Put $f_1(x) \equiv xg_1'(x)$. Then $f_1'(x) = g_1'(x) + xg_1''(x) \leq g_1'(x)$. So since f_1 and g_1 agree on $[-1, 1]$, it follows that $|f_1(x)| \leq |g_1(x)|$ for all x . Put $g_n(x) \equiv ng_1(x/n)$ and $f_n(x) \equiv xg_1'(x/n) = xg_1'(x/n)$. Of course then $|f_n(x)| \leq |g_n(x)|$ for all x .

Now consider $h_n(x) \equiv g_n(x^2)$. By (iii) we have (now $x \in \mathbb{R}^n$ again.)

$$\frac{d}{dt} \int h_n(x) \rho(x, t) dx = \int v(x, t) \cdot \nabla h_n(x) \rho(x, t) dx.$$

But $\nabla h_n(x) = 2xg_1'(x^2)$, so $|\nabla h_n(x)|^2 = 4x^2g_1''(x^2) \cdot g_1'(x^2) = 4f_n(x^2) \cdot g_1'(x^2)$. Clearly g_1' takes its values in $[0, a]$ for some a independent of n , and we are free to assume $a \leq 1$. Then we have $|\nabla h_n(x)|^2 \leq 4f_n(x^2) \leq 4g_n(x^2) = 4h_n(x)$. By the Schwarz inequality,

$$\frac{d}{dt} \int h_n(x) \rho(x, t) dx \leq 2 \left(\int (v^2(x, t) \rho(x, t) dx) \right)^{1/2} \cdot \left(\int h_n(x) \rho(x, t) dx \right)^{1/2}. \quad (7)$$

Let $m_n(t) = \left(\int h_n(x) \rho(x, t) dx \right)$. Then $(d/dt)(m_n(t))^{1/2} = \frac{1}{2}(m_n(t))^{-1/2} (d/dt)m_n(t) \leq N$, where $N \equiv \sup_{t \in [0, T]} \left(\int v^2(x, t) \rho(x, t) dx \right)^{1/2}$, which is finite by (iii). We then have $m_n(t)^{1/2} \leq m_n(0)^{1/2} + Nt$, so $m_n(t) \leq 2 \left(\left(\int x^2 \rho(x, 0) dx \right) + N^2 t^2 \right)$. The right-hand side is independent of n , and since $h_n(x)$ increases monotonically to x^2 , an application of the monotone convergence theorem implies:

$$\int x^2 \rho(x, t) dx \leq \int 2x^2 \rho(x, 0) dx + \left(2 \sup_{s \leq t} \int v^2(x, s) \rho(x, s) dx \right) t^2. \quad (8)$$

This concludes the proof.

Results similar to (iv) can be found in [16]. Inequality (8) is a sharper statement of (iv) than seems to be available. We will use it in Sect. IV.

III. Solution of the Backward Martingale Equation

Let u, v and ρ be given as in Theorem 2.1, and fix a compact time interval $[0, T]$. The backward martingale equation for u and v is:

$$\frac{\partial f}{\partial t}(x, t) = \left(\frac{1}{2} \Delta - (v - u) \cdot \nabla \right) f(x, t), \quad f(x, t) = f_0(x). \quad (1)$$

In this section we will solve (1) for $f_0 \in L^2(\rho(x, 0) dx)$. The solutions will be weak solutions, but they will have enough regularity for our construction. As the first step, consider a sequence of smooth bounded \mathbb{R}^n valued functions $b_*^i(x, t)$ on $\mathbb{R}^n \times [0, T]$. (The superscript i no longer tells which component of a vector field is considered, as in Sect. I, but instead which vector field.) These are to approximate the singular function $v - u$ which will be b_* for the diffusion we construct in the

next section. We will choose the sequence so that

$$\lim_{i \rightarrow \infty} \int_0^T \int |\dot{b}_*^i - (v - u)|^2 \rho(x, t) dx dt = 0, \tag{2}$$

which is seen possible by an easy density argument. We assume in the following that such a choice has been fixed. Now choose an initial condition f_0 . At this time we require some smoothness of f_0 , so take $f_0 \in C_0(\mathbb{R}^n)$. By $f^i(x, t)$ we will mean the solution to the initial value problem:

$$\frac{\partial f^i}{\partial t}(x, t) = (\frac{1}{2}\Delta - b_*^i \cdot \nabla) f^i(x, t), \quad f^i(x, 0) = f_0(x). \tag{3}$$

Since b_*^i is smooth and bounded and f_0 is in the domain of $(\frac{1}{2}\Delta - b_*^i \cdot \nabla)$ (in the natural function space), existence, uniqueness and regularity of the solution to (3) are easy to establish and well known. In particular, we have the maximum principle for solutions of (3):

$$\min_x f_0(x) \leq f^i(y, t) \leq \max_x f_0(x) \quad \text{for all } (y, t) \in \mathbb{R}^n \times [0, T]. \tag{4}$$

We will make heavy use of this fact.

In some sense (3) approximates (1), so one might hope that the solutions of (3) approximate the solutions of (1) and thus guarantee their existence. This hope is borne out in the following results.

Theorem 3.1 *Let f_0, b_*^i and f^i be as above. Then*

$$\lim_{i, j \rightarrow \infty} \int (f^i - f^j)^2 \rho(x, t) dx = 0, \text{ converging uniformly in } t,$$

$$\lim_{i, j \rightarrow \infty} \int_0^T \int |\nabla f^i - \nabla f^j|^2 \rho(x, t) dx dt = 0. \tag{5}$$

Proof. We will establish three inequalities from which our result follows. Let $c \equiv \max |f_0(x)|$. Wherever c appears, we have used the maximum principle (4). To obtain our inequalities, we repeatedly integrate by parts. We begin with some preparatory comments.

First, by our assumptions on b_*^i and f_0, f^i is bounded and has continuous bounded derivatives of all orders. Recall that where $\rho \neq 0, u(x, t) = \frac{1}{2}\nabla \rho(x, t) / \rho(x, t)$ ((12) of Sect. I). We define \mathcal{D} to be the space of continuous functions on \mathbb{R}^n with bounded first and second continuous derivatives. Then for any g and h in \mathcal{D} , the following integration by parts is valid:

$$\int (\nabla g \cdot \nabla h) \rho(x, t) dx = - \int 2g(\frac{1}{2}\Delta + u \cdot \nabla) h \rho(x, t) dx \text{ a.e. } dt. \tag{6}$$

The operator on the right, $-(\frac{1}{2}\Delta + u(x, t) \cdot \nabla)$ is evidently positive on $L^2(\rho(x, t) dx)$. We denote its Friedrichs extension by O_t and call it the *osmotic operator*. It will sometimes be convenient to write $(\cdot, \cdot)_t$ for the inner product on $L^2(\rho(x, t) dx)$ and $\|\cdot\|_t$ for the norm. We can then rewrite (6) as

$$\int (\nabla g \cdot \nabla h) \rho(x, t) dx = 2(g, O_t h)_t \text{ a.e. } dt. \tag{7}$$

Second, f^i will have continuous bounded time derivatives for all $t \in [0, T]$ —this is where we need f_0 smooth. Suppose $g(x, t)$ is a function on $\mathbb{R}^n \times [0, T]$ such that each $g(\cdot, t)$ is in \mathcal{D} and $(\partial g / \partial t)$ is continuous. Then the weak continuity equation implies:

$$-2 \int_0^T \int \left(g \frac{\partial g}{\partial t} \right) \rho(x, t) dx dt = 2 \int_0^T \int g(v \cdot g) \rho(x, t) dx dt + \|g\|_0^2 - \|g\|_T^2. \tag{8}$$

Now we compute. Using (6), (3) and (8) and the fact that f^i satisfies (3) we obtain:

$$\begin{aligned} \int_0^T \int \nabla f^i \cdot \nabla f^i \rho(x, t) dx dt &= -2 \int_0^T \int f^i \left(\frac{1}{2} \Delta + u \cdot \nabla \right) f^i \rho(x, t) dx dt \\ &= -2 \int_0^T \int f^i \left(\left(\frac{\partial}{\partial t} + b_*^i \cdot \nabla + u \cdot \nabla \right) f^i \right) \rho(x, t) dx dt \\ &= -2 \int_0^T \int f^i \left((-v \cdot \nabla + b_*^i \cdot \nabla + u \cdot \nabla) f^i \right) \rho(x, t) dx dt + \|f^i\|_0^2 - \|f^i\|_T^2 \\ &\leq 2c \left(\int_0^T \int |b_*^i - (v - u)|^2 \rho dx dt \right)^{1/2} \left(\int_0^T \int |\nabla f^i|^2 \rho dx dt \right)^{1/2} + 2c^2. \end{aligned}$$

Applying the quadratic formula to this quadratic inequality, we get:

$$\begin{aligned} \left(\int_0^T \int |\nabla f^i|^2 \rho dx dt \right)^{1/2} &\leq c \left(\int_0^T \int |b_*^i - (v - u)|^2 \rho dx dt \right)^{1/2} \\ &\quad + c \left(\int_0^T \int |b_*^i - (v - u)|^2 \rho dx dt + 4 \right)^{1/2}. \end{aligned} \tag{9}$$

we can now prove the second part of (5). Let $w \equiv f^i - f^j$. Then:

$$\begin{aligned} \int_0^T \int \nabla w \cdot \nabla w \rho dx dt &= -2 \int_0^T \int w \left(\frac{1}{2} \Delta + u \cdot \nabla \right) w \rho dx dt \\ &= -2 \int_0^T \int w \left(\frac{\partial f^i}{\partial t} - \frac{\partial f^j}{\partial t} + (u + b_*^i) \cdot \nabla f^i - (u + b_*^j) \cdot \nabla f^j \right) \rho dx dt \\ &= -2 \int_0^T \int w \left(\frac{\partial w}{\partial t} + (u + b_*^i) \cdot \nabla f^i - (u + b_*^j) \cdot \nabla f^j \right) \rho dx dt \\ &= 2 \int_0^T \int w (b_*^j - (v - u)) \cdot \nabla f^j - (b_*^i - (u - v)) \cdot \nabla f^i \rho dx dt \\ &\quad + \|w\|_0^2 - \|w\|_T^2 \\ &\leq 4c \left(\left(\int_0^T \int |b_*^i - (v - u)|^2 \rho dx dt \right)^{1/2} \left(\int_0^T \int |\nabla f^i|^2 \rho dx dt \right)^{1/2} \right. \\ &\quad \left. + \left(\int_0^T \int |b_*^j - (v - u)|^2 \rho dx dt \right)^{1/2} \left(\int_0^T \int |\nabla f^j|^2 \rho dx dt \right)^{1/2} \right), \end{aligned} \tag{10}$$

where we have used the fact that $w(x, 0) \equiv 0$.

To prove the first part of (5), we need to estimate $\|w\|_t^2$. Clearly:

$$\begin{aligned} \int w^2 \rho(x, t) dx &= \int_0^t \frac{d}{ds} \int w^2 \rho(x, s) dx ds, \text{ and } \frac{d}{ds} \int w^2 \rho(x, s) dx \\ &= 2 \int w \frac{\partial w}{\partial s} \rho(x, s) dx + 2 \int w(v \cdot \nabla w) \rho(x, s) dx. \end{aligned}$$

So since

$$\begin{aligned} \frac{\partial w}{\partial s} &= \frac{\partial f^i}{\partial s} - \frac{\partial f^j}{\partial s} = \frac{1}{2} \Delta w + b_*^j \cdot \nabla f^j - b_*^i \cdot \nabla f^i \\ &= \frac{1}{2} (\Delta w + (b_*^j - b_*^i) \cdot \nabla (f^i + f^j) - (b_*^i + b_*^j) \cdot \nabla w), \end{aligned}$$

and since $\int w(\Delta w) \rho dx = - \int \nabla w \cdot \nabla w \rho dx - 2 \int w(u \cdot \nabla w) \rho dx$, we have:

$$\begin{aligned} \frac{d}{ds} \int w^2 \rho dx &= - \int \nabla w \cdot \nabla w \rho dx - 2 \int w(u \cdot \nabla w) \rho dx + \int w(b_*^j - b_*^i) \cdot \nabla (f^i + f^j) \rho dx \\ &\quad - \int w(b_*^i + b_*^j) \cdot \nabla w dx + 2 \int w(v \cdot \nabla w) \rho dx \\ &= - \int \nabla w \cdot \nabla w \rho dx + \int w(b_*^j - b_*^i) \cdot \nabla (f^i + f^j) \rho dx \\ &\quad - \int w(b_*^i + b_*^j - 2(v - u)) \cdot \nabla w \rho dx. \end{aligned}$$

From this last equation, we deduce the inequality:

$$\begin{aligned} \left(\frac{d}{ds} \int w^2 dx \right)_+ &\leq 2c \{ (\int |b_*^i - b_*^j|^2 \rho dx)^{1/2} ((\int |\nabla f^i|^2 \rho dx)^{1/2} + (\int |\nabla f^j|^2 \rho dx)^{1/2}) \\ &\quad + (\int |b_*^i + b_*^j - 2(v - u)|^2 \rho dx)^{1/2} (\int |\nabla w|^2 \rho dx)^{1/2} \}. \end{aligned}$$

Integrating the left-hand side up to time t and the right-hand side up to time T we get:

$$\begin{aligned} \int (f^i - f^j)^2 \rho(x, t) dx &\leq 2c \left\{ \left(\int_0^T \int |b_*^i - b_*^j|^2 \rho dx ds \right)^{1/2} \left(\left(\int_0^T \int |\nabla f^i|^2 \rho dx ds \right)^{1/2} \right. \right. \\ &\quad \left. \left. + \left(\int_0^T \int |\nabla f^j|^2 \rho dx ds \right)^{1/2} \right) + \left(\int_0^T \int |b_*^i + b_*^j - 2(v - u)|^2 \rho dx ds \right)^{1/2} \right. \\ &\quad \left. \cdot \left(\int_0^T \int |\nabla w|^2 \rho dx ds \right)^{1/2} \right\}. \end{aligned} \tag{11}$$

Taking (9), (10) and (11) together, we get the rest of (5).

Henceforth let $f(x, t)$ denote the limit of our approximating sequence $f^i(x, t)$, which exists by the above theorem. For each $t, f(\cdot, t)$ is well defined as an element of $L^2(\rho(x, t))$; we denote it by f_t .

Corollary 3.1. $\min_x f_0(x) \leq f(x, t) \leq \max_x f_0(x)$ a.e. $\rho(x, t) dx$ for each t . (12)

Proof. This is an immediate consequence of the fact that

$$\left\{ g : \min_x f_0(x) \leq g(x) \leq \max_x f_0(x) \right\}$$

is a closed subset of $L^2(\rho(x, t)dx)$ and (4).

Theorem 3.2. *Let f and f_0 be as above. Then :*

$$\int_s^t \int |\nabla f(x, \tau)|^2 \rho(x, \tau) dx d\tau + \|f_t\|_t^2 = \|f_s\|_s^2, \tag{13}$$

$$\int f(x, t)\rho(x, t)dx = \int f(x, s)\rho(x, s)dx, \tag{14}$$

for each $0 \leq s \leq t \leq T$. Furthermore, we have the estimate :

$$\|f_t^i\|_t \leq k_c^i + \|f_s^i\|_s, \tag{15}$$

where $f_t^i(x) \equiv f^i(x, t)$ and k_c^i is a constant depending only on b_*^i and c with the property that $\lim_{i \rightarrow \infty} k_c^i = 0$.

Proof.

$$\begin{aligned} \frac{1}{2} \int_0^T \frac{d}{d\tau} |f_\tau^i|^2 \rho dx d\tau &= \int_0^T \left(\left(f_\tau^i, \frac{\partial f_\tau^i}{\partial \tau} \right)_\tau + (v \cdot \nabla f_\tau^i, f_\tau^i) \right) d\tau \\ &= \int_0^T \left((f_\tau^i, (\frac{1}{2}\Delta - b_*^i \cdot \nabla) f_\tau^i)_\tau + (v \cdot \nabla f_\tau^i, f_\tau^i) \right) d\tau \\ &= \int_0^T \left((f_\tau^i, (\frac{1}{2}\Delta + u \cdot \nabla) f_\tau^i)_\tau - ((b_*^i - (v - u)) \cdot \nabla f_\tau^i, f_\tau^i)_\tau \right) d\tau \\ &= -\frac{1}{2} \int_0^T (\nabla f_\tau^i, \nabla f_\tau^i) d\tau - \int_0^T ((b_*^i - (v - u)) \cdot \nabla f_\tau^i, f_\tau^i) d\tau. \end{aligned}$$

Let $k_c^i \equiv 2c \left(\int_0^T \int |b_*^i - (v - u)|^2 dt \right)^{1/2} \cdot (\text{r.h.s. of (9)})$; this has the required properties (see inequality (9)). Then the above computation gives us $\|f_t^i\|_t^2 \leq \|f_s^i\|_s^2 + k_c^i - \int_s^t (\nabla f_\tau^i, \nabla f_\tau^i) d\tau$, which is even stronger than (15). Since by (11) and (9) $\lim_{i \rightarrow \infty} \|f_t^i\|_t^2 = \|f_t\|_t^2$ (and the same for s) and $\lim_{i \rightarrow \infty} \int_s^t (\nabla f_\tau^i \cdot \nabla f_\tau^i) \rho dx d\tau = \int_s^t (\nabla f_\tau \cdot \nabla f_\tau) \rho dx d\tau$, (13) follows from the above.

Next compute $\int_s^t \left(\frac{d}{d\tau} \int f^i \rho(x, \tau) dx \right) d\tau$.

$$\frac{d}{d\tau} \int f^i \rho(x, \tau) dx = \int (\frac{1}{2}\Delta + u \cdot \nabla) f^i \rho(x, \tau) dx - \int (b_*^i - (v - u)) \cdot \nabla f^i \rho(x, \tau) dx.$$

The term on the left is zero a.e. dt since it equals $(1, O_\tau f^i)_\tau = (O_\tau 1, f^i)_\tau = 0$ for a.e. τ . (Clearly $1 \in \mathcal{D}(O_\tau)$, and $O_\tau 1 = 0$.)

The term on the right vanishes as our approximation is tightened:

$$\int f^i \rho(x, t) dx - \int f^i \rho(x, s) dx \leq \left(\int_s^t \int (b_*^i - (v - u))^2 \rho dx d\tau \right)^{1/2} \left(\int_s^t |\nabla f^i|^2 \rho dx d\tau \right)^{1/2}.$$

This concludes the proof since L^2 convergence implies L^1 convergence for a probability space.

Equation (13) allows us to relax our restrictions on the initial condition f_0 . Let $P_{t,0}$ be the operator from $C_0^\infty(\mathbb{R}^n)$ to $L^2(\rho(x, t) dx)$ given by $P_{t,0} : f_0 \rightarrow f_t$. By (13) and the corollary to Theorem 3.1, $P_{t,0}$ extends by continuity to a Markovian contraction, also denoted $P_{t,0}$,

$$P_{t,0} : L^2(\rho(x, t) dx) \rightarrow L^2(\rho(x, t) dx).$$

(To say that $P_{t,0}$ is Markovian means $P_{t,0}$ is positivity preserving and $P_{t,0} 1 = 1$.) Note that while at each time t we consider a different Hilbert space, each of them contains all the bounded measurable functions.

There is nothing special about $t = 0$. We can carry out the exact same construction for any interval $[s, T]$, $s < T$. Let $P_{t,s} : L^2(\rho(x, s) dx) \rightarrow L^2(\rho(x, t) dx)$ be the operator so constructed; clearly it too is a Markovian contraction. Our next result says that taken together, the family of operators $\{P_{t,s}\}_{s < t}$ form a *Markovian propagator*:

Theorem 3.3. *Let $r \leq s \leq t$. Then $P_{t,r} = P_{t,s} P_{s,r}$.*

Proof. This is a consequence of the last theorem and the trivial but important fact that our solution f is independent of the particular approximating sequence b_*^i . Let b_{*1}^i be an approximating sequence for $(v - u)$ on $[r, s]$, and let b_{*2}^i be an approximating sequence for $(v - u)$ on $[s, t]$. Define b_*^i by $b_*^i = b_{*1}^i$ for times in $[r, s]$, and $b_*^i = b_{*2}^i$ for times in $[s, t]$. Then b_*^i is an approximating sequence for $(v - u)$ on $[r, t]$; clearly the possible discontinuity at time s is immaterial. Choose an $f_s \in C_0^\infty(\mathbb{R}^n)$ and again put $c = \max_x |f(x)|$. Let $P_{b,a}^i$ be the propagator for the approximating equation ((3) with b_*^i as above) so that $P_{a,s}^i f_s(x) = f^i(x, a)$, $s < a$, gives the approximating solutions. Then we have:

$$\begin{aligned} \|P_{t,r} f_r - P_{t,s} P_{s,r} f_r\|_t &\leq \|P_{t,r} f_r - P_{t,r}^i f_r\|_t + \|P_{t,r}^i f_r - P_{t,s} P_{s,r} f_r\|_t \\ &\leq \|P_{t,r} f_r - P_{t,r}^i f_r\|_t + \|P_{t,r}^i f_r - P_{t,s} P_{s,r}^j f_s\|_t + \|P_{t,s} (P_{s,r} f_r - P_{s,r}^j f_s)\|_t. \end{aligned}$$

The first and third terms vanish as i and j go to infinity by Theorem 3.1. The second term is trickier. By uniqueness for the approximating Cauchy problem, we have $P_{t,r}^i = P_{t,s}^i P_{s,r}^i$, so this term can be rewritten as:

$$\|P_{t,s}^i P_{s,r}^i f_r - P_{t,s} P_{s,r}^j f_r\|_t \leq \|P_{t,s}^i (P_{s,r}^i - P_{s,r}^j) f_r\|_t + \|(P_{t,s} - P_{t,s}^i) P_{s,r}^j f_r\|_t.$$

By (15), the first term on the right is dominated by:

$$k_{2c}^i + \|P_{s,r}^i f_r - P_{s,r}^j f_r\|_s.$$

This too vanishes as i and j go to infinity. By (9), (10) and (11), we see that $(P_{t,s} - P_{t,s}^i)$

goes to zero uniformly on sets of the form $\{f : a \leq f(x) \leq b\}$ for any a and b in \mathbb{R} . Since $P_{s,r}^j$ is Markovian, we are almost done. It only remains to observe that $C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\rho(x, t)dx)$.

We now discuss the sense in which f solves the backward martingale equation. In general, f will only have first spatial derivatives and no time derivative, so we can only expect f to be a weak solution. We have the following result:

Theorem 3.4. *Let $f \in L^2(\rho(x, s)dx)$ and $g \in \mathcal{D}$. Then :*

$$(P_{t,s}f, g)_t - (f, g)_s = \int_s^t ((P_{\tau,s}f, O_\tau g)_\tau + (P_{\tau,s}f, v \cdot \nabla g)_\tau) d\tau. \tag{16}$$

When $f \in \mathcal{D}$ also, $\tau \mapsto (P_{\tau,s}f, g)$ is right differentiable at $\tau = s$ with derivative $(f, O_s g)_s + (f, v \cdot \nabla g)_s$.

Proof. First take $f, g \in \mathcal{D}$, and integrating by parts on our approximate solutions, obtain (16) in this case. Then approximate an arbitrary $f \in L^2(\rho(x, s)dx)$ with a sequence $f_n \in \mathcal{D}$. Since (16) holds for each f_n and g , and since $(P_{\tau,s}f_n, O_\tau g)_\tau + (P_{\tau,s}f_n, v \cdot \nabla g)_\tau$ converges to $(P_{\tau,s}f, O_\tau g)_\tau + (P_{\tau,s}f, v \cdot \nabla g)_\tau$ uniformly in τ , we get the first part.

Smoothness of f can be used in the following way: using (16), (13) and the weak continuity equation, it is easy to establish that for $f \in \mathcal{D}$ $\lim_{h \rightarrow 0^+} \|P_{t+h,t}f - f\|_{t+h} = 0$.

Then since $(f, v \cdot \nabla g)_t = \text{Im} \langle \nabla \psi_t, (\nabla g) f \psi_t \rangle$ and $(f, u \cdot \nabla g)_t = \text{Re} \langle \nabla \psi_t, (\nabla g) f \psi_t \rangle$ (where $\langle \cdot, \cdot \rangle$ denotes the inner product on $L^2(\mathbb{R}^n)$), it is easy to see that the integrand in (16) is continuous at $\tau = s$. Knowing this, we can divide both sides of (16) by h and take the $h \rightarrow 0^+$ limit to complete the proof.

Equation (16) may not look much like (1), but if we write $f(x, t)$ for $P_{t,s}f(x)$, the left-hand side *formally* equals

$$\begin{aligned} \int \frac{\partial f}{\partial t}(x, t)g(x)\rho(x, t)dx + \int (v(x, t) \cdot \nabla f(x, t))g(x)\rho(x, t)dx \\ + \int (v(x, t) \cdot \nabla g(x))f(x, t)\rho(x, t)dx, \end{aligned}$$

and the right-hand side *formally* equals:

$$\int ((\frac{1}{2}A + u(x, t) \cdot \nabla) f(x, t))g(x)\rho(x, t)dx + \int f(x, t)(v(x, t) \cdot \nabla g(x))\rho(x, t)dx.$$

Cancelling and rearranging, we would then have:

$$\left(\frac{\partial f}{\partial t} - \frac{1}{2}A f + (v - u) \cdot \nabla f \right) g \rho(x, t) dx = 0 \text{ for all } g \text{ in } C_0(\mathbb{R}^n),$$

so (16) is the right result.

There are many further questions one can ask about the backward martingale equation; however, the above theorems give us all the information we require in the next section, and so we end our investigation here.

The method of combining the maximum principle and action estimates ($\int \nabla f \cdot \nabla f \rho dx$ is usually called an energy integral, so it's natural to call

$\int_0^T \int \nabla f \cdot \nabla f \rho dx dt$ an action integral) has been employed by Nelson in [17] in a different context. The method can only be applied to second order equations where one has the maximum principle; but when it can be applied, it is at least as powerful as the more usual coercivity methods; e.g. those in Lions [18]. There is no coercivity estimate for our equation; the $v \cdot \nabla$ term is just too singular.

IV. Construction of the Diffusion

We assume given an interval $[0, T]$ and functions u, v and ρ as in the last section. We make the convention that wherever r, s and t appear denoting elements of \mathbb{R} , $0 \leq r \leq s \leq t \leq T$. Let $P_{t,s}$ denote the Markovian propagator of Theorem 3.3.

By *path space* Ω , We mean $(\mathbb{R}^n)^{[0,T]}$ with the product topology. (The dot indicates one point compactification.) Let $\xi(t)$ also denote the t -configuration function on Ω , defined just as $\xi(t)$ is for trajectory space. The Borel field of \mathcal{B} is then just $\mathcal{B} = \sigma\{\xi(t): t \in [0, T]\}$, which is of the same form as the Borel field for Ω_c . (The topologies bear no such relation.) The past, present and future are defined and denoted as before. Clearly $\Omega_c \subset \Omega$; and it is a theorem of Nelson that Ω_c is an $F_{\sigma\delta}$ in Ω , and hence a measurable set. In this section we will first employ a method of Nelson to construct a probability measure Pr on Ω in terms of $P_{t,s}$ and $\rho(x, T)$. It will then be easy to show that under Pr , $\xi(t)$ is a square integrable Markov process such that $\xi(t)$ has density $\rho(x, t)$. To complete our analysis, we will show that

$$t \mapsto w(t) \equiv \xi(t) - \xi(0) - \int_0^t b(\xi(\tau), \tau) d\tau \tag{1}$$

defines a Wiener process under Pr . This will imply that $Pr(\Omega_c) = 1$, and that the restriction of Pr to Ω_c is the unique (regular) Borel measure on Ω_c making $\xi(t)$ a diffusion with density ρ and forward drift b . By b we mean $u + v$, and by b_* we will mean $v - u$.

To construct the measure on Ω , consider $\mathcal{C}_{sp}(\Omega)$, the set of all continuous functions F on Ω which can be written as a finite sum of finite products of functions of the form $\omega \mapsto f(\xi(t, \omega))$ for some $f \in C^\infty(\mathbb{R}^n)$. Such a representation is never unique; we can always include a trivial dependence on $\xi(t)$ for any t . This will not cause any problems, and in fact we take advantage of it by always assuming an $\xi(T)$ dependence in every product we write. This will simplify some notation. Now each $f \in C^\infty(\mathbb{R}^n)$ is identified with a multiplication operator on $L^2(\rho(x, t) dx)$ for each t in a natural way. We can then define a linear functional $Pr : \mathcal{C}_{sp}(\Omega) \rightarrow \mathbb{R}$ by:

$$Pr(F) = (f_n, P_{T,t_{n-1}} f_{n-1} P_{t_{n-1},t_{n-2}} f_{n-2} \cdots P_{t_2,t_1} f_1)_T \tag{2}$$

for F of the form $F(\omega) = \prod_{i=1}^n f_i(\xi(t_i, \omega))$ with each $f_i \in C^\infty(\mathbb{R}^n)$ and $t_n = T$, and then taking the linear extension to $\mathcal{C}_{sp}(\Omega)$. By Theorem 3.3, the right-hand side is independent of the particular representation of F , and $F \mapsto Pr(F)$ is a well defined positive linear functional on $\mathcal{C}_{sp}(\Omega)$ with $Pr(1) = 1$. Since $\mathcal{C}_{sp}(\Omega)$ is an algebra containing 1 and separating points, it is uniformly dense in $\mathcal{C}(\Omega)$. Pr therefore

extends by continuity to a positive linear functional on $\mathcal{C}(\Omega)$, also denoted Pr , with $Pr(1) = 1$. Then by the Riesz–Markov theorem, there is a unique regular Borel probability measure, again denoted Pr , so that:

$$Pr(F) = \int F(\omega)Pr(d\omega) = E(F). \tag{3}$$

This method of constructing measures on function spaces from a Markovian propagator differs from the usual Kolmogorov extension method in the important feature that it gives us a regular Borel measure on a compact Hausdorff space. This greatly facilitates the investigation of essentially topological questions such as sample path continuity. The basic idea is due to Kakutani; it was developed into a method by Nelson in [19].

Let Pr denote the measure just constructed; all other notation will be as before. Assume henceforth that $\int x^2\rho(x, 0)dx < \infty$.

Lemma 4.1. $t \mapsto \xi(t)$ is a square integrable Markov process under Pr which has a jointly measurable version. Moreover :

- (i) The image of Pr under $\xi(t)$ has density $\rho(x, t)$ with respect to Lebesgue measure
- (ii) $t \mapsto \xi(t)$ is strongly continuous into $L^2(Pr)$.
- (iii) For any f and g in \mathcal{D} (as in Sect. III), the following limits hold :

$$\lim_{h \rightarrow 0^+} \frac{1}{h} E((f(\xi(t)) - f(\xi(t-h)))g(\xi(t))) = -((\frac{1}{2}\Delta - (v-u) \cdot \nabla)f, g)_t, \tag{4}$$

$$\lim_{h \rightarrow 0^+} \frac{1}{h} E((f(\xi(t+h)) - f(\xi(t)))g(\xi(t))) = (\frac{1}{2}\Delta + (u+v) \cdot \nabla)f, g)_t. \tag{5}$$

Proof. First we note that by the construction of Pr , $\xi(t)$ is clearly Markovian. Next we verify (i). Let $g \in C_0^\infty(\mathbb{R}^n)$. Then by (2) and (3):

$$\begin{aligned} E(g(\xi(t))) &= (1, P_{T,t}g)_T \\ &= \int \rho(x, T)P_{T,t}g(x)dx \\ &= \int \rho(x, t)g(x)dx \end{aligned}$$

by (14) of Theorem 3.2. This proves the assertion about the density of $\xi(t)$. With this in hand we know that $E(\xi^2(t)) = \int x^2\rho(x, t)dx$, and so by virtue of (iv) of Theorem 2.1, $t \rightarrow E(\xi^2(t))$ is a bounded measurable function on $[0, T]$. In fact, $E((\xi(t) - \xi(s))^2) = E(\xi^2(t)) + E(\xi^2(s)) - 2E(\xi(t)\xi(s))$. An easy argument of the type used to prove part (iv) of Theorem 2.1 together with Theorem 3.2 shows that $|E(\xi^2(s)) - E(\xi(t)\xi(s))| \leq k(t-s)$, where k depends only on the energy. Since $|\int x^2\rho(x, t)dx - \int x^2\rho(x, s)dx| \leq h(t-s)$ for some constant h (as in (iv) of Theorem 2.1), we then have the strong continuity claimed in (i).

Now the Bochner–Von Neumann theorem doesn't apply directly to $\xi(t)$; but by the strong continuity established above, the $\xi(t)$ lie in a separable subspace of $L^2(Pr)$, and so the argument used to prove it applies. We therefore conclude that our process has a jointly measurable version. It remains to prove (iii). We again

compute using (2) and (3):

$$\begin{aligned} & \frac{1}{h} E((f(\xi(t)) - f(\xi(t-h)))g(\xi(t))) \\ &= \frac{1}{h} ((f, g)_t - (P_{t,t-h}f, g)_t) \\ &= \frac{1}{h} ((f, g)_t - (f, g)_{t-h}) + \frac{1}{h} ((f, g)_{t-h} - (P_{t,t-h}f, g)_t). \end{aligned}$$

Now apply the weak continuity equation to the first term on the bottom line and Theorem 3.4 to the second. Taking the limit $h \rightarrow 0^+$ we get:

$$((v \cdot \nabla f, g)_t + (f, v \cdot \nabla g)_t) + (\frac{1}{2}(\nabla f, \nabla g)_t - (f, v \cdot \nabla g)_t).$$

Cancelling and integrating by parts, we obtain (4). Next:

$$\begin{aligned} & \frac{1}{h} E(((f(\xi(t+h)) - f(\xi(t)))g(\xi(t)))) \\ &= \frac{1}{h} ((f, P_{t+h,t}g)_{t+h} - (f, g)_t) \\ &= \frac{1}{h} \int_t^{t+h} ((0_\tau f, P_{\tau,t}g)_\tau + (v \cdot \nabla f, P_{\tau,t}g)_\tau) d\tau \end{aligned}$$

as in the proof of Theorem 3.4. So, for f and g in \mathcal{D} , it's clear that taking the $h \rightarrow 0^+$ limit we get, by continuity of the integrand, $(0_t f, g)_t + (v \cdot \nabla f, g)_t$. This proves (5).

Implicit in the above is the fact that $E(b^2(\xi(t), t)) = \int (u + v)^2 \rho(x, t) dx$ and this is uniformly bounded on $[0, T]$ by Theorem 2.1. Moreover, it's clear $(t, \omega) \mapsto b(\xi(t, \omega), t)$ has a jointly measurable version which is square integrable on $[0, T] \times \Omega$. Therefore, by the Fubini theorem we can integrate pathwise to form $\int_s^t b(\xi(\tau), \tau) d\tau$, with a harmless abuse of notation. This random variable is clearly \mathcal{P}_t and \mathcal{F}_s measurable and in $L^2(Pr)$. For future use we record the following estimate of its $L^2(Pr)$ norm

Lemma 4.2. $E\left(\left(\int_s^t b(\xi(\tau), \tau) d\tau\right)^2\right) \leq k(t-s)^2$, where k is independent of t and s .

Proof. $E\left(\int_s^t b(\xi(\tau), \tau) d\tau\right)^2 \leq E\left(n(t-s)^{1/2} \left(\int_s^t b^2(\xi(\tau), \tau) d\tau\right)^{1/2}\right)^2$
 $= n^2(t-s) \int_s^t E b^2(\xi(\tau), \tau) d\tau \leq n^2(t-s)h$ for some constant h by

Theorem 2.1 Now we define $\eta(t, s)$ by:

$$\eta(t, s) = \xi(t) - \xi(s) - \int_s^t b(\xi(\tau), \tau) d\tau. \tag{7}$$

Clearly $\eta(t, s)$ is \mathcal{P}_t and \mathcal{F}_s measurable and in $L^2(Pr)$. Intuitively, we are subtracting the drift and initial value out of our diffusion on $[s, t]$ in forming $\eta(t, s)$. Since the

second order part of our generator is $\frac{1}{2}\Delta$, $\eta(t, s)$ should then be the increment of a Brownian motion on $[s, t]$. As indicated, we will prove this. We begin with:

Lemma 4.3. *For each $y \in \mathbb{R}^n$ and each $s \in [0, T]$, let $Z_y^s(t) \equiv \exp(iy \cdot \eta(t, s) + y^2(t - s)/2)$. Then $EZ_y^s(t) = 1$ for all t .*

Proof. First notice that $Z_y^s(t)Z_y^t(t + h) = Z_y^s(t + h)$. Also, for each ω we have:

$$Z_y^s(t + h) = Z_y^s(t) \exp(y^2 h/2) \exp(-iy \cdot \xi(t)) \exp(iy \cdot \xi(t + h)) \\ - \int_t^{t+h} Z_y^s(t) \exp(iy \cdot (\xi(t + h) - \xi(t) + y^2 h/2 - \int_t^\tau b(\xi(\sigma), \sigma) d\sigma)) iy \cdot b(\xi(\tau), \tau) d\tau.$$

Now take the expectation of the term on the right. By a simple Taylor expansion argument using the estimate (7), the error in replacing the exponential factor in the integrand with 1 is $o(h)$ uniformly in t . (By Fubini's theorem orders of integration can be interchanged.) So, the expectation of the term on the right is:

$$- \int_t^{t+h} E(Z_y^s(t) iy \cdot b(\xi(\tau), \tau)) d\tau + o(h)$$

with the error uniform in t .

Next, write $E_t(Z_y^s(t)) = g(\xi(t))$; g will be (complex) bounded. Then using the definition (2) and Theorem 3.4:

$$E(Z_y^s(t) \exp(-iy \cdot \xi(t)) \exp(iy \cdot \xi(t + h))) \\ = E(g(\xi(t)) \exp(iy \cdot \xi(t)) \exp(iy \cdot \xi(t + h))) \\ = (g, 1)_t + \int_t^{t+h} (P_{\tau,t} f^* g, (\frac{1}{2}\Delta + b \cdot \nabla) f)_\tau d\tau,$$

where $f(x) \equiv \exp(iy \cdot x)$ and f^* is its complex conjugate. (The above inner product notation is slightly abusive. We are working with real Hilbert spaces throughout, and no complex conjugation is taken in forming the implied integrals.) Since $(\frac{1}{2}\Delta + b \cdot \nabla) f = (-y^2/2 + iy \cdot b) f$, we can rewrite the last integral as:

$$\int_t^{t+h} (P_{\tau,t} f^* g, (-y^2/2 + iy \cdot b) f)_\tau d\tau \\ = \int_t^{t+h} E(Z_y^s(t) \exp(iy \cdot (\xi(\tau) - \xi(t))) (-y^2/2 + iy \cdot b)) \\ = \int_t^{t+h} E(Z_y^s(t) (-y^2/2 + iy \cdot b) + o(h)),$$

where as above, the error term is uniform in t .

Again, an easy Taylor series argument shows that $t \rightarrow EZ_y^s(t)$ is continuous, and $EZ_y^s(t) = 1 + O(t - s)$. Finally, since $\exp(y^2 h/2) = 1 + y^2 h/2 + o(h)$, we have altogether that $EZ_y^s(t + h) - EZ_y^s(t) = o(h)$ uniformly in t , and so the lemma is proved.

Corollary 4.1. $\eta(t, s)$ is gaussian with mean 0 and covariance $\delta^{ij}(t - s)$.

Proof. The characteristic function of $\eta(t, s)$ is $E(\exp(iy \cdot \eta(t, s)))$, and by the above, this is $\exp(-y^2(t - s)/2)$.

Lemma 4.4. $t \mapsto w(t) \equiv \eta(t, s)$ is an $(\Omega, \mathcal{P}_t, Pr)$ Brownian motion started at s .

Proof. The adaptation is obvious, and with the above corollary, the connection with the heat equation is immediate. The lemma is perhaps therefore trivial, but I know of no theorem that can be quoted to assert the path continuity. However, with the above information on the increments, Nelson’s proof of path continuity for the Wiener process [20] applies to our process with one modification: The sets $\{\omega : |w(t_i) - w(t_j)| > 2\epsilon\}$ are here known only to be measurable instead of open. But since Pr is regular, they can be approximated on the outside by open sets with an error of less than $2\rho(\epsilon/2, \delta)$ (using the notation of [20]). Using these sets, the rest of the proof goes through with a factor of 2, and so $t \mapsto w(t)$ is continuous with probability 1.

At last we come to our existence theorem.

Theorem 4.1. Given u, v and ρ as above, there exists a Borel probability measure \tilde{Pr} on Ω_c such that:

- (i) $t \mapsto \xi(t)$ is a square integrable Markov process under \tilde{Pr} which has a jointly measurable version.
- (ii) The image of \tilde{Pr} under $\xi(t)$ has density $\rho(x, t)$.
- (iii) $w(t) \equiv \xi(t) - \xi(0) - \int_0^t b(\xi(\tau), \tau) d\tau$ is an $(\Omega_c, \mathcal{P}_t, \tilde{Pr})$ Brownian motion.
- (iv) For any $f \in C_0^\infty(\mathbb{R}^n)$, the following limits exist strongly in $L^2(\tilde{Pr})$:

$$\lim_{h \rightarrow 0^+} \frac{1}{h} E_t(f(\xi(t+h)) - f(\xi(t))) = (\frac{1}{2}\Delta + b(\xi(t), t) \cdot \nabla) f(\xi(t)),$$

$$\lim_{h \rightarrow 0^+} \frac{1}{h} E_t(f(\xi(t)) - f(\xi(t-h))) = -(\frac{1}{2}\Delta - b_*(\xi(t), t) \cdot \nabla) f(\xi(t)).$$

Proof. By the Fubini theorem, $t \mapsto \int_0^t b(\xi(\tau), \tau) d\tau$ is almost surely continuous. By Lemma 4.4, $t \mapsto \eta(t, 0)$ is almost surely continuous. Therefore, $t \mapsto \xi(t)$ is almost surely continuous, and $Pr(\Omega_c) = 1$. Let \tilde{Pr} be the restriction of Pr to Ω_c . Then (i) and (ii) are immediate consequences of Lemma 4.1, and (iii) is an immediate consequence of Lemma 4.4. To prove (iv), employ a Taylor expansion, using obvious estimates on the Wiener terms and Lemma 4.2 on the drift terms.

The configuration process under \tilde{Pr} is therefore our desired diffusion. (A diffusion is sometimes required to be a family of probability measures $\tilde{Pr}_{x,s}$ on Ω_c under which $\tilde{Pr}_{x,s}\{\xi(s) = x\} = 1$, as in [11]. These can be recovered from \tilde{Pr} since Ω_c is a Polish space and has regular conditional probabilities.)

As for uniqueness, we make the following comment. Consider an approximating sequence of backward drifts as in Sect. III. For our given final density, there is a unique measure Pr^i making the configuration process a diffusion with final density $\rho(x, T)$ and backward drift $b_*^i(x, t)$. It is not hard to show \tilde{Pr} as above is the weak

limit of the Pr^i . Clearly, there is only one solution to our problem with this stability property.

Conclusion.

The results presented here provide the solutions of the stochastic differential equations

$$d\xi(t) = b(\xi(t), t)dt + dw(t), \quad (1)$$

arising in stochastic mechanics when the configuration space is \mathbb{R}^n and the potential is Rellich class. The methods are of considerably greater generality than the results. For instance, if the configuration space is a compact Riemannian manifold, the PDE results of Sect. III and the construction of the measure in Sect. IV go through virtually unchanged. Therefore one can take $\mathbb{R}^3 \times \text{SO}(3)$ as the configuration space and following Dankel [21] incorporate spin into stochastic mechanics. It is also easy to generalize the class of admissible interactions so that magnetic fields are included using virtually the same analysis as that in Sect. II. Other generalizations are quite likely possible. Finally we observe that the time interval on which we worked was arbitrary, and so we have constructed a probability measure on the trajectory space for all times.

The project of developing stochastic mechanics into a detailed mathematical theory has only recently begun, and it is impossible to predict the scope of the theory to emerge. Nonetheless, it seems unlikely that detailed knowledge of the sample paths of these conservative diffusions would fail to shed light on mathematical problems arising in ordinary quantum mechanics, at the very least.

For example, Shucker [22] has shown that, under some assumptions on the initial conditions, if the solutions to (1) exist for zero potential (i.e. when b comes from a solution of the free Schroedinger equation), then

$$\lim_{t \rightarrow \infty} \frac{1}{t} \xi(t, \omega) \equiv p(\omega) \text{ exists a.s.} \quad (2)$$

and that p has the same distribution as the ordinary quantum mechanical momentum. I have recently proved such a result for a more natural class of initial conditions and for a wide class of potentials. These results will appear in a stochastic mechanical treatment of quantum scattering.

Acknowledgements. I would like to thank my teacher Edward Nelson for many helpful and enlightening conversations. In particular, I would like to thank him for encouraging me to use PDE methods to attack the problem. I would also like to thank him and John Lafferty for a critical reading of the first draft of this paper

References

1. Nelson, E.: Quantum Pluctuations. Princeton: Princeton University Press 1984
2. Guerra, F., Morato, L.: Quantization of dynamical systems and stochastic control theory. Phys. Rev. **D27**, 1774–1786 (1983)

3. Fényes, I.: Eine wahrscheinlichkeitstheoretische Begründung und Interpretation der Quantenmechanik. *Z. Phys.* **132**, 81–106 (1952)
4. Nelson, E.: Derivation of the Schrödinger equation from Newtonian mechanics. *Phys. Rev.* **150**, 1079–1085 (1966)
5. Yasue, K.: Stochastic calculus of variations. *J. Funct. Anal.* **41**, 327–340 (1981)
6. Zambrini, J.: Preprint (to appear)
7. Zheng, W., Meyer, P.: Quelques resultats de “Mecanique Stochastique.” Preprint to appear in *Seminaire de Probabilités XVII*, J. Azéma ed., Berlin, Heidelberg. New York: Springer 1984
8. Albeverio, S., Høegh-Krohn, R.: A remark on the connection between stochastic mechanics and the heat equation. *J. Math. Phys.* **15**, 1745–1747 (1974)
9. Carmona, R.: Processus de diffusion gouverné par la form de Dirichlet de l’opérateur de Schrödinger. In *Seminaire de Probabilités XIII*, C Dellacherie, Meyer, P. A., Weil, M. (eds.). Berlin, Heidelberg, New York: Springer 1979
10. Nelson, E.: Critical diffusions. Preprint to appear in *Seminaire de Probabilités XVIII*, Azéma. J. (ed.). Berlin Heidelberg, New York, Tokyo Springer 1985
11. Strook, D., Varadhan, S.: *Multidimensional Diffusion Processes*. Berlin, Heidelberg, New York: Springer 1979
12. Reed, M., Simon, B.: *Methods of modern mathematical physics*, Vol. 2, New York: Academic Press 1975
13. Bochner, S., Von Neumann, J.: On compact solutions of operator differential equations I. *Ann. Math. Ser. 2*, **36**, 255–291 esp. 263 (1935)
14. Doob, J.: Stochastic processes depending on a continuous parameter. *Trans. Am. Math. Soc.* **42**, 107–140 (1937)
15. Kato, T.: Some inequalities for linear operators. *Math. Ann.* **125**, 208–212 (1952)
16. Radin, C., Simon, B.: Invariant domains for the time dependent Schrödinger equation. *J. Diff. Eqn.* **29**, 289–296 (1978)
17. Nelson, E.: Les écoulements incompressibles d’énergie finie, *Colloques internationaux du Centre national de la recherche scientifique No 117, Les équations aux dérivées partielles*, Editions du C.N.R.S., Paris, 1962.
18. Lions, J.: *Equations différentielles operationelles et problèmes aux limits*. Berlin, Heidelberg, New York: Springer 1961
19. Nelson, E.: Regular probability measures on function space. *Ann. Math.* **69**, 630–643 (1959)
20. Nelson, E.: Feynman integrals and the Schrödinger equation. *J. Math. Phys.* **5**, 332–343, esp 339 (1964)
21. Dankel, T.: Mechanics on manifolds and the incorporation of spin into Nelson’s stochastic mechanics. *Arch. Rat. Mech. Anal.* **37**, 192–222 (1970)
22. Shucker, D.: Stochastic mechanics of systems with zero potential. *J. Funct. Anal.* **38**, 146–155 (1980)

Communicated by A. Jaffe

Received January 20, 1984

Note added in proof. The proof given of Lemma 4.4 is brief enough that it borders on the cryptic. In particular, it should explain why $\eta(t, s)$ is independent of \mathcal{P}_s . Professor P.A. Meyer kindly pointed this out in a letter to the author and offered the following explanation (slightly rephrased): Let $A \in \mathcal{R}^n$ be such that $Pr\{\xi(s) \in A\} = 1/\alpha > 0$. Let $\tilde{Z}_y^s(t) = \alpha \chi(\xi(s)) Z_y^s(t)$, where $Z_y^s(t)$ is given as in lemma 4.3. Then the statement and proof of lemma 4.3 hold with $Z_y^s(t)$ replaced by $\tilde{Z}_y^s(t)$. This shows that conditional on $\xi(s) \in A$, $\eta(t, s)$ is still mean zero with covariance $\delta^{ij}(t - s)$. Together with the Markov property, this shows that $\eta(t, s)$ is independent of \mathcal{P}_s .

Also, murky phrases such as “close to being a small perturbation” occurring in the introduction do have precise, but much lengthlier, formulations. These can be found in the author’s Princeton thesis.

