

Variational Processes from the Weak Forward Equation

Toshio Mikami

Department of Mathematics, Tokyo Institute of Technology, Ohokayama, Meguro-ku, Tokyo 152, Japan

Received September 11, 1989; in revised form June 7, 1990

Abstract. In this paper the author constructs Markov diffusion processes from a given system of Borel probability measures on a d -dimensional Euclidean space. He constructs a, so-called, variational process which does not always coincide with a Nelson process. He also discusses Schrödinger’s problem in quantum mechanics.

0. Introduction

The theory of Markov processes has been developed by many authors (cf. Dynkin [10, 11]). But

(Q) Under what information can we assume that the real world is Markovian?

Of course, it depends on how we consider the real world. But it is better if we can assume that it is Markovian, because Markov property is a nice (kind of ideal) property. Before we state our problem, we mention that some notations are given in the end of this section.

In this paper we consider the following problem; let us fix $T > 0$. Assume that we are given the system of Borel probability measures $\{\rho(t, dx)\}_{0 \leq t \leq T}$ on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ which satisfies the following weak forward equation; for any $f \in C_b^{1,2}([0, T] \times \mathfrak{R}^d; \mathfrak{R})$ and any $0 \leq s \leq t \leq T$,

$$\begin{aligned}
 & \int_{\mathfrak{R}^d} f(t, x)\rho(t, dx) - \int_{\mathfrak{R}^d} f(s, x)\rho(s, dx) \\
 &= \int_s^t \int_{\mathfrak{R}^d} [\partial f(u, x)/\partial u + \left(\sum_{i,j=1}^d a^{ij}(u, x)\partial^2 f(u, x)/\partial x_i \partial x_j \right)] / 2 \\
 &+ \sum_{i=1}^d b^i(u, x)\partial f(u, x)/\partial x_i] \rho(u, dx) du,
 \end{aligned} \tag{0.1}$$

where $a(t, x) \equiv (a^{ij}(t, x))_{i,j=1}^d$ is a Borel measurable, positive definite, symmetric $d \times d$ -matrix and $b(t, x) \equiv (b^i(t, x))_{i=1}^d$ is a Borel measurable d -dimensional vector. In our setting, $a(t, x)$ and $b(t, x)$ can depend on $\rho(t, dx)$ implicitly (cf. Application 2 below) and the following conditions are crucial;

$$\int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x), b(t, x) \rangle \rho(t, dx) dt < \infty, \quad (0.2)$$

$$\rho(t, dx)/dx \equiv \rho(t, x) \text{ exists for all } 0 < t \leq T. \quad (0.3)$$

Problem. *Is there a Markov diffusion process $\{X(t)\}_{0 \leq t \leq T}$ such that for some Borel measurable $\tilde{b}(\cdot, \cdot)$ from $[0, T] \times \mathfrak{R}^d$ to \mathfrak{R}^d ,*

$$dX(t) = \tilde{b}(t, X(t))dt + a(t, X(t))^{1/2}dW(t), \quad (0.4)$$

$$P(X(t) \in dx) = \rho(t, dx) \text{ for all } 0 \leq t \leq T, \quad (0.5)$$

where $W(t)$ is a d -dimensional $\sigma[X(s); s \leq t]$ -Wiener process?

The answer to the above problem is YES if $a(t, x)$ is nice (see Sect. 1, Theorem 1.1).

Remark 0.1. As we mention in Sect. 3, $\tilde{b}(t, x)$ cannot always be determined uniquely from $\{\rho(t, dx)\}_{0 \leq t \leq T}$. Otherwise, all the properties of the path space (i.e. $X(\cdot)$) turn out to be determined by the one-dimensional marginal distributions.

Definition 0.1. *We call $b(t, x)$ the mean forward derivatives of $\{\rho(t, dx)\}_{0 \leq t \leq T}$ (cf. Nelson [30, 31]).*

Remark 0.2. If the answer to the problem is YES, then (0.2) is a finiteness condition of the relative entropy of $X(\cdot)$ with respect to $X^0(\cdot)$ which satisfies the following stochastic differential equation in a weak sense (cf. Stroock and Varadhan [37]);

$$dX^0(t) = a(t, X^0(t))^{1/2}dW^0(t), \quad (0.6)$$

$$P(X^0(0) \in dx) = \rho(0, dx), \quad (0.7)$$

where $W^0(t)$ is a d -dimensional $\sigma[X^0(s); s \leq t]$ -Wiener process (cf. Sect. 3, the proof of Theorem 1.1).

Definition 0.2. *We call the condition (0.2) the finite energy condition.*

For what kind of problems can we apply the above result?

Applications are the following;

(Application 1). When we only consider one-dimensional marginal distributions of continuous semimartingales (here we assume that martingale parts are square integrable), we do not have to consider non-Markovian semimartingales; let $X(t)$ be a semimartingale on some probability space $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{0 \leq t \leq T}, P)$ which satisfies

$$dX(t) = \beta(t, \omega)dt + dM(t), \quad (0.8)$$

where $\beta(t, \omega) = (\beta^i(t, \omega))_{i=1}^d$ is (\mathfrak{F}_t) -progressively measurable and satisfies

$$\int_0^T \int_{\Omega} |\beta(t, \omega)| dt P(d\omega) < +\infty, \quad (0.9)$$

and $M(t) = (M(t)^i)_{i=1}^d$ is a square integrable (\mathfrak{F}_t) -martingale starting from 0 whose

quadratic variational processes are absolutely continuous with respect to dt , P -a.s. (cf. Ikeda and Watanabe [20]). Denote quadratic variational processes by the following:

$$d\langle M^i, M^j \rangle(t, \omega) = \alpha^{ij}(t, \omega) dt \quad \text{for all } i, j = 1, \dots, d, \quad (0.10)$$

where we take an (\mathfrak{F}_t) -progressively measurable version $\alpha^{ij}(t, \omega)$, which is possible from the integrability conditions of $\langle M^i, M^j \rangle(t, \omega)$ ($i, j = 1, \dots, d$).

Then by the Ito formula (cf. Ikeda and Watanabe [20]), for any $f \in C_b^{1,2}([0, T] \times \mathfrak{R}^d; \mathfrak{R})$ and any $0 \leq s \leq t \leq T$,

$$\begin{aligned} & E[f(t, X(t))] - E[f(s, X(s))] \\ &= \int_s^t E \left[\partial f(v, X(v)) / \partial v + \left(\sum_{i,j=1}^d \alpha^{ij}(v, \omega) \partial^2 f(v, X(v)) / \partial x_i \partial x_j \right) \right] / 2 \\ & \quad + \sum_{i=1}^d \beta^i(v, \omega) \partial f(v, X(v)) / \partial x_i \Big] dv \\ &= \int_s^t E \left[\partial f(v, X(v)) / \partial v + \left(\sum_{i,j=1}^d \tilde{\alpha}^{ij}(v, X(v)) \partial^2 f(v, X(v)) / \partial x_i \partial x_j \right) \right] / 2 \\ & \quad + \sum_{i=1}^d \tilde{\beta}^i(v, X(v)) \partial f(v, X(v)) / \partial x_i \Big] dv, \end{aligned} \quad (0.11)$$

where we denote by E the mean value with respect to P and put

$$\tilde{\alpha}^{ij}(t, x) = E[\alpha^{ij}(t, \omega) | (t, X(t) = x)], \quad (0.12)$$

$$\tilde{\beta}^i(t, x) = E[\beta^i(t, \omega) | (t, X(t) = x)]. \quad (0.13)$$

Here the mean value is that by conditioning $X(t) = x$ (cf. Nelson [31]). Therefore if the finite energy condition (0.2) is satisfied and $\tilde{\alpha}(t, x)$ is nice, then we can construct a Markov process which satisfies (0.4) and (0.5) with $a(t, x) \equiv \tilde{\alpha}(t, x)$ (cf. Sect. 1, Theorem 1.1). Hence as far as we consider problems on semimartingales which are related only to their one-dimensional marginal distributions, we can replace semimartingales by Markov diffusion processes (cf. Mikami [28], Sect. 4 for more details).

(Application 2). This Problem is quite similar to that of constructing a Nelson process. Assume that we are given a normalized wave function $\psi(t, x)$ (a solution of Schrödinger equation); on $[0, T] \times \mathfrak{R}^d$,

$$\begin{aligned} & (-1)^{1/2} \partial \psi(t, x) / \partial t = -\Delta_x \psi(t, x) / 2 + V(x) \psi(t, x), \\ & \int_{\mathfrak{R}^d} |\psi(t, x)|^2 dx = 1 \quad \text{for all } 0 \leq t \leq T, \end{aligned} \quad (0.14)$$

for some function $V(\cdot): \mathfrak{R}^d \rightarrow \mathfrak{R}$ (cf. Carlen [3], Sakurai [33]).

Then $|\psi(t, x)|^2 dx$ satisfies (0.1) with $a(t, x) \equiv$ an identity matrix and

$$b(t, x) = \begin{cases} \operatorname{Re}(\nabla_x \psi(t, x) / \psi(t, x)) + \operatorname{Im}(\nabla_x \psi(t, x) / \psi(t, x)) & \text{if } \psi(t, x) \neq 0; \\ 0 & \text{if } \psi(t, x) = 0. \end{cases} \quad (0.15)$$

If $V(\cdot)$ is a Rellich class potential, then there exists a Markov process which satisfies (0.4) and (0.5) (cf. Carlen [3], [4]).

(Application 3). Schrödinger's problem. We can also consider the following problem (cf. Schrödinger [34, 35]); let us assume that we are given systems of Borel probability measures $\{r(t, dx)\}_{0 \leq t \leq T}$ whose end point distributions $r(0, dx)$ and $r(T, dx)$ are independent of all r . Then what kind of r minimizes the energy integral (0.2)? The answer is a system of one-dimensional marginal distributions of a Markovian Bernstein process ($P(\phi)_1$ -process). For more details, see Sect. 4.

In Sect. 1, we define and prove the existence and the uniqueness of a Variational Process which is the unique Markov diffusion process whose family of one-dimensional marginal distributions is a given solution $\{\rho(t, dx)\}_{0 \leq t \leq T}$ of a weak forward equation and whose mean forward derivative (cf. Nelson [30, 31]) attains the minimum of the energy integral among those with which $\{\rho(t, dx)\}_{0 \leq t \leq T}$ satisfies the weak forward equation.

In Sect. 2 we give some results necessary for the proof of the existence and the uniqueness of a variational process.

In Sect. 3 we prove the existence and the uniqueness of a variational process. There we mention that the mean forward derivatives cannot always be determined uniquely from a given $\{\rho(t, dx)\}_{0 \leq t \leq T}$. In the proof, Markovian Bernstein processes ($P(\phi)_1$ -processes), a dynamic programming equation, Fleming's logarithmic transformation, Föllmer's entropy approach and his results on the time reversal of Markov processes play crucial roles.

In Sect. 4, we discuss Schrödinger's problem.

Finally we give some notations.

- $\mathfrak{R}^d \equiv$ a d -dimensional Euclidean space.
- $\mathfrak{R} \equiv \mathfrak{R}^1$.
- $o \equiv$ the origin in \mathfrak{R}^d .
- $|\beta| \equiv \left(\sum_{i=1}^d |\beta^i|^2 \right)^{1/2}$ for $\beta = (\beta^i)_{i=1}^d \in \mathfrak{R}^d$.
- $\langle \cdot, \cdot \rangle \equiv$ an inner product in \mathfrak{R}^d .
- $\text{Id} \equiv$ a $d \times d$ -identity matrix.
- $\nabla_x \equiv (\partial/\partial x_i)_{i=1}^d$.
- $\Delta_x \equiv \sum_{i=1}^d \partial^2/\partial x_i^2$.
- $B(S) \equiv$ the smallest σ -field which contains all open subset of a topological space S .
- $X(t) \equiv X(t, \omega)$.
- $\sigma[X(s) | s \leq t] \equiv$ the smallest σ -field which makes the map $\omega \mapsto X(s, \omega)$ measurable for all $0 \leq s \leq t$.
- $C_b^{1,2}([0, T] \times \mathfrak{R}^d, \mathfrak{R}) \equiv$ the set of functions $f(t, x)$ of $[0, T] \times \mathfrak{R}^d$ to \mathfrak{R} with continuous and bounded derivatives up to the first order with respect to t and up to the second order with respect to x , respectively.
- $C(A; B) \equiv$ a set of continuous functions from a set A to a set B whose topology is that induced by the sup norm.
- $C_o^\infty(A; B) \equiv$ a set of infinitely differentiable functions from a set A to a set B with compact supports.
- $C_b^\infty(A; B) \equiv$ a set of functions from a set A to a set B with bounded derivatives of all orders.

1. Variational Processes

In this section we consider the weak forward equation by way of a variational method and introduce variational processes. From now on, we fix $T > 0$.

Let us define the variational problem associated with a solution $\{\rho(t, dx)\}_{0 \leq t \leq T}$ of the weak forward equation.

Definition 1.1. For a given solution $\{\rho(t, dx)\}_{0 \leq t \leq T}$ of the weak forward equation (0.1), let us denote by $\mathbf{A} = \mathbf{A}_\rho$ the set of Borel measurable functions $b(\cdot, \cdot) = \{b^i(\cdot, \cdot)\}_{i=1}^d$ of $[0, T] \times \mathfrak{R}^d$ to \mathfrak{R}^d such that for any $f \in C_b^{1,2}([0, T] \times \mathfrak{R}^d; \mathfrak{R})$ and any $0 \leq s \leq t \leq T$,

$$\begin{aligned} & \int_{\mathfrak{R}^d} f(t, x) \rho(t, dx) - \int_{\mathfrak{R}^d} f(s, x) \rho(s, dx) \\ &= \int_s^t \int_{\mathfrak{R}^d} \left[\partial f(u, x) / \partial u + \left(\sum_{i,j=1}^d a^{ij}(u, x) \partial^2 f(u, x) / \partial x_i \partial x_j \right) / 2 \right. \\ & \quad \left. + \sum_{i=1}^d b^i(u, x) \partial f(u, x) / \partial x_i \right] \rho(u, dx) du. \end{aligned} \quad (1.1)$$

The variational problem associated with $\{\rho(t, dx)\}_{0 \leq t \leq T}$ is to find the optimal $b^o \in \mathbf{A}$ such that

$$\begin{aligned} & \inf_{b^o \in \mathbf{A}} \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)', b(t, x)' \rangle \rho(t, dx) dt / 2 \\ &= \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^o, b(t, x)^o \rangle \rho(t, dx) dt / 2 < +\infty. \end{aligned} \quad (1.2)$$

Let us state assumptions.

(A.0). $\rho(t, dx)/dx \equiv \rho(t, x)$ exists for all $0 < t \leq T$.

(A.1). The infimum in (1.2) is finite.

(A.2). $a(t, x) \equiv (a^{ij}(t, x))_{i,j=1}^d$ is a uniformly positive definite, symmetric $d \times d$ -matrix on $[0, T] \times \mathfrak{R}^d$ which is bounded, once continuously differentiable and uniformly Hölder continuous. $\nabla_x a(t, x)$ is bounded and the first derivatives of $a(t, x)$ are uniformly Hölder continuous with respect to x .

Then we get the following theorem which is proved in Sect. 3.

Theorem 1.1. Suppose that (A.0)–(A.2) hold. Then for a solution $\{\rho(t, dx)\}_{0 \leq t \leq T}$ of the weak forward equation (0.1), there exist a unique solution b^o of the variational problem associated with $\{\rho(t, dx)\}_{0 \leq t \leq T}$ and a unique \mathfrak{R}^d -valued Markov process $\{X(t)\}_{0 \leq t \leq T}$ on $(C([0, T]; \mathfrak{R}^d), B(C([0, T]; \mathfrak{R}^d)))$ such that

$$dX(t) = b(t, X(t))^o dt + a(t, X(t))^{1/2} dW(t), \quad (1.3)$$

where $W(t)$ is a d -dimensional $\sigma[X(s); s \leq t]$ -Wiener process and that

$$P(X(t) \in dx) = \rho(t, dx) \quad \text{for all } 0 \leq t \leq T \quad (1.4)$$

and that for almost all $u \in (0, T]$, $\rho(u, x)$ is absolutely continuous and for any $0 < t \leq T$,

$$\int_t^T \int_{\mathfrak{R}^d} |b(s, x)^o - \nabla_x \rho(s, x) / \rho(s, x)|^2 \rho(s, dx) ds < \infty. \quad (1.5)$$

Remark 1.1. Suppose that (A.2) holds. Then (A.0) and (A.1) are also necessary conditions in Theorem 1.1. (A.1) holds automatically. (A.0) can be proved in the following way; $\rho(t, dx) = P(X(t) \in dx)$ is absolutely continuous with respect to $P(X^0(t) \in dx)$ (cf. Remark 0.2 for notation) for all $0 \leq t \leq T$ from (A.1) (cf. Liptser and Shiryaev [24], Theorem 7.10) and $P(X^0(t) \in dx)$ is absolutely continuous with respect to dx for all $0 < t \leq T$ from (A.2) (cf. Stroock and Varadhan [37]).

Definition 1.2. We call the Markov process $\{X(t)\}_{0 \leq t \leq T}$ in Theorem 1.1 a variational process associated with a solution $\{\rho(t, dx)\}_{0 \leq t \leq T}$ of the weak forward equation (0.1). We also call b^o the optimal mean forward derivative (cf. Nelson [30, 31]).

Remark 1.2. There are other results concerning this problem. Carlen's result [3] whose generalization can be found in Carlen [4] is by way of a semigroup approach (see also Krylov [23]). Its probabilistic counterpart is Carmona [5]. Nagasawa [29] is by way of the transformations of measures and their setting is different from Carlen's. Zheng [39] is by way of the convergence of semimartingales.

Remark 1.3. The definition of a variational process is different from Zambrini's [38]. In [38], Zambrini considered, by the dynamic programming equation, the variational problem with constraints on marginal distributions at time $t = 0$ and $t = T$ and only semimartingales (see also Dai Pra [8]). Föllmer's approach is direct, simple and does not have to use the dynamic programming equation (cf. Föllmer [14], Sect. 2.1.3). In our case, the constraints are on one-dimensional marginal distributions in the whole time interval $[0, T]$ and we consider systems of Borel probability measures on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$, which includes those of one-dimensional marginal distributions of semimartingales (cf. Sect. 0, Application 1). And the dynamic programming equation plays a crucial role (we cannot avoid using it). The problem considered by Zambrini was originally considered by Schrödinger (cf. [34, 35]). We call the problem Schrödinger's Problem and extend the class of one-dimensional marginal distributions under considerations from that of semimartingales to that of Borel probability measures on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ in Sect. 4.

2. Lemmas

In this section we give results necessary for the proof of Theorem 1.1. All the assumptions can be found in Sect. 1. First we introduce Bernstein processes (one-dimensional local Markov random fields), especially Markovian Bernstein processes ($P(\phi)_1$ -processes). A Bernstein process is also called a reciprocal process. The original idea of Bernstein process can be found in E. Schrödinger's articles [34, 35], where he tried to consider quantum mechanics by way of the probability theory. E. Schrödinger's idea was studied, independently, by S. Bernstein [1].

In this section the state space of the stochastic processes is \mathfrak{R}^d .

Let us give some definitions.

Definition 2.1. For $T > 0$, let $\{X(t)\}_{0 \leq t \leq T}$ be an \mathfrak{R}^d -valued stochastic process on some probability space. We call $\{X(t)\}_{0 \leq t \leq T}$ a Bernstein process (one-dimensional local Markov random field) iff for any $0 \leq s < t < u \leq T$ and any $A \in B(\mathfrak{R}^d)$,

$$P(X(t) \in A | X(v), v \in [0, s] \cup [u, T]) = P(X(t) \in A | X(s), X(u)). \quad (2.1)$$

Definition 2.2. (cf. Jamison [21]). *The functions $\{p(s, x; t, y; u, z)\}_{0 \leq s < t < u \leq T, x, y, z \in \mathfrak{R}^d}$ are called Bernstein transition probability density functions iff*

(1) *for any $0 \leq s < t < u \leq T$ and $x, z \in \mathfrak{R}^d$, the map*

$$B(\mathfrak{R}^d) \in A \mapsto \int_A p(s, x; t, y; u, z) dy \quad (2.2)$$

defines a probability measure on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$,

(2) *for any $0 \leq s < t < u \leq T$ and any $A \in B(\mathfrak{B}^d)$, the map*

$$(x, z) \mapsto \int_A p(s, x; t, y; u, z) dy \quad (2.3)$$

is $B(\mathfrak{R}^d) \times B(\mathfrak{R}^d)$ -measurable, and

(3) *for any $0 \leq s < t < u < v \leq T$, $x, z \in \mathfrak{R}^d$ and any $A, C \in B(\mathfrak{R}^d)$,*

$$\begin{aligned} & \int_A p(s, x; u, y; v, z) \left(\int_C p(s, x; t, w; u, y) dw \right) dy \\ &= \int_C p(s, x; t, y; v, z) \left(\int_A p(t, y; u, w; v, z) dw \right) dy. \end{aligned} \quad (2.4)$$

Let Ω be the set of \mathfrak{R}^d -valued functions on $[0, T]$ (this notation has nothing to do with that in Sect. 0).

Let Σ be the smallest σ -field such that the map

$$\pi_t: \Omega \mapsto \mathfrak{R}^d, \quad \pi_t(\omega) \equiv \omega(t) \quad (2.5)$$

is measurable for all $0 \leq t \leq T$.

Let $\{q(s, x; t, y)\}_{0 \leq s < t \leq T, x, y \in \mathfrak{R}^d}$ be continuous, strictly positive transition probability density functions of $X^0(\cdot)$ (cf. Remark 0.2), where we assume (A.2) (cf. Stroock and Varadhan [37]).

For any $0 \leq s < t < u \leq T$ and $x, y, z \in \mathfrak{R}^d$, put

$$p(s, x; t, y; u, z) = q(s, x; t, y)q(t, y; u, z)/q(s, x; u, z). \quad (2.6)$$

Then $\{p(s, x; t, y; u, z)\}_{0 \leq s < t < u \leq T, x, y, z \in \mathfrak{R}^d}$ become Bernstein transition probability density functions (cf. Definition 2.2) and for each probability measure μ on $B(\mathfrak{R}^d) \times B(\mathfrak{R}^d)$, there exists a unique \mathfrak{R}^d -valued Bernstein process (one-dimensional local Markov random field) $\{X(t) \equiv X^\mu(t)\}_{0 \leq t \leq T}$ on (Ω, Σ) such that for any $0 \leq s < t < u \leq T$,

$$\begin{aligned} P(X(t) \in dz | X(s), X(u)) &= p(s, X(s); t, z; u, X(u)) dz, \\ P(X(0) \in dx, X(T) \in dy) &= \mu(dx dy) \end{aligned} \quad (2.7)$$

(cf. Jamison [21], Theorem 2.1).

The following theorem characterizes a Markovian Bernstein process $(P(\phi))_1$ -process) constructed from a Markov process as above.

Theorem 2.1. (cf. Jamison [21], Theorem 3.1). *Let $\{X(t)\}_{0 \leq t \leq T}$ be a Bernstein process (one-dimensional local Markov random field) on (Ω, Σ) constructed from $\{q(s, x; t, y)\}_{0 \leq s < t \leq T, x, y \in \mathfrak{R}^d}$ and μ as above. Then $\{X(t)\}_{0 \leq t \leq T}$ is a Markov process iff there exist σ -finite measures ν_0 and ν_T on $B(\mathfrak{R}^d)$ such that*

$$\mu(dx dy) = q(0, x; T, y) \nu_0(dx) \nu_T(dy). \quad (2.8)$$

Remark 2.1. The Markov process constructed in Theorem 2.1 is a h -path process by Doob (cf. Doob [9]) and is called a Markovian Bernstein process ($P(\phi)_1$ -process) constructed from $q(s, x; t, y)$ and μ (see also Simon [36]).

Next let us consider Schrödinger's functional equations which determine the σ -finite measures ν_0 and ν_T on $B(\mathfrak{R}^d)$ for one-dimensional marginal distributions of a given probability measure μ on $B(\mathfrak{R}^d) \times B(\mathfrak{R}^d)$ in Theorem 2.1.

Definition 2.3. For probability measures μ_0 and μ_T on $B(\mathfrak{R}^d)$ and $q(0, x; T, y)$, the following equations are called Schrödinger's functional equations;

$$\mu_0(dx) = \nu_0(dx) \int_{\mathfrak{R}^d} q(0, x; T, y) \nu_T(dy), \quad (2.9)$$

$$\mu_T(dy) = \nu_T(dy) \int_{\mathfrak{R}^d} q(0, x; T, y) \nu_0(dx), \quad (2.10)$$

where ν_0 and ν_T are σ -finite measures on $B(\mathfrak{R}^d)$.

Schrödinger's functional equations have been considered by Beurling [2], Fortet [15] and Jamison [21]. The following theorem together with Theorem 2.1 imply the existence and the uniqueness of a Markovian Bernstein process ($P(\phi)_1$ -process) constructed from two end point distributions $\mu_0(dx)$, $\mu_T(dx)$ and $q(s, x; t, y)$.

Theorem 2.2. (cf. Jamison [21], Theorem 3.2.). For given probability measures μ_0 and μ_T on $B(\mathfrak{R}^d)$, there exists a unique solution (ν_0, ν_T) of Schrödinger's functional equations.

From now on, we assume that we are given end point distributions $\mu_0(dx)$, $\mu_T(dx)$ and the solutions $\nu_0(dx)$, $\nu_T(dx)$ of Schrödinger's functional equations for $\mu_0(dx)$, $\mu_T(dx)$ (cf. Definition 2.3 and Theorem 2.2).

Put, for $\nu_T(dy)$ in (2.10),

$$h(t, x) \equiv \int_{\mathfrak{R}^d} q(t, x; T, y) \nu_T(dy). \quad (2.11)$$

Then the following results are known.

Lemma 2.3. (cf. Jamison [22], p. 330.). Suppose that (A.2) holds. Then $h(t, x) \in C^{1,2}([0, T] \times \mathfrak{R}^d; \mathfrak{R})$ and on $(0, T) \times \mathfrak{R}^d$,

$$\partial h(t, x) / \partial t = - \left(\sum_{i,j=1}^d a^{ij}(t, x) \partial^2 h(t, x) / \partial x_i \partial x_j \right) / 2. \quad (2.12)$$

Theorem 2.4. (cf. Jamison [22], Theorem 2 or Dai Pra [8], Theorem 2.1.). Suppose that (A.2) holds. If $\nu_T(dy)$ in (2.10) is absolutely continuous with respect to the Lebesgue measure dy , then the Markovian Bernstein process ($P(\phi)_1$ -process) constructed as above is a weak solution (cf. Stroock and Varadhan [37]) of

$$dX(t) = a(t, X(t)) \nabla_x \log h(t, X(t)) dt + a(t, X(t))^{1/2} dW(t) \quad \text{on } [0, T]. \quad (2.13)$$

where $W(t)$ is a d -dimensional $\sigma[X(s); s \leq t]$ -Wiener process.

Before we give a crucial lemma, let us give some notations and a definition. Put

$$\mathcal{Q}_T \equiv \{q_1, q_2, \dots\}: \text{the set of rational numbers in } (0, T). \quad (2.14)$$

Take

$$\mathcal{Q}_{n,T} \equiv \{q_0^n \equiv 0, q_1^n, \dots, q_{n-1}^n, q_n^n \equiv T\} \quad \text{for all } n \geq 1, \quad (2.15)$$

so that

$$\mathcal{Q}_{n,T} \subset \mathcal{Q}_{n+1,T} \quad \text{for all } n \geq 1, \quad (2.16)$$

$$\bigcup_{n=2}^{\infty} \mathcal{Q}_{n,T} \setminus \{0, T\} = \mathcal{Q}_T. \quad (2.17)$$

Definition 2.4. A system of Borel probability measures $\{r(t, dx)\}_{0 \leq t \leq T}$ is in the set \mathbf{A}_n iff there exists a Borel measurable function $b(t, x)^r$ from $[0, T] \times \mathfrak{R}^d$ to \mathfrak{R}^d such that for any $f \in C_b^{1,2}([0, T] \times \mathfrak{R}^d; \mathfrak{R})$ and any $0 \leq s \leq t \leq T$,

$$\begin{aligned} & \int_{\mathfrak{R}^d} f(t, x)r(t, dx) - \int_{\mathfrak{R}^d} f(s, x)r(s, dx) \\ &= \int_s^t \int_{\mathfrak{R}^d} [\partial f(u, x)/\partial u + \left(\sum_{i,j=1}^d a^{ij}(u, x) \partial^2 f(u, x) / \partial x_i \partial x_j \right)] / 2 \\ & \quad + \sum_{i=1}^d b^i(u, x)^r \partial f(u, x) / \partial x_i] r(u, dx) du, \end{aligned} \quad (2.18)$$

$$r(q_i^n, dx) = \rho(q_i^n, dx) \quad \text{for all } i = 0, \dots, n \quad (2.19)$$

(cf. (2.15)).

Put

$$L_n \equiv \inf \left\{ \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^r, b(t, x)^r \rangle r(t, dx) dt / 2; r \in \mathbf{A}_n, b^r \right\}. \quad (2.20)$$

Here we take the infimum over all possible b^r for each $r \in \mathbf{A}_n$.

Remark 2.2. L_n is bounded from (A.1) and (A.2), since $\rho \in \mathbf{A}_n$ for all $n \geq 1$.

Define $b(t, x)^n$, piecewise on $[q_i^n, q_{i+1}^n) (i = 0, \dots, n - 1)$, by,

$$b(t, x)^n \equiv a(t, x) \nabla_x \log h_i^n(t, x) \quad (2.21)$$

(cf. Theorem 2.4), where we put

$$h_i^n(t, x) = \int_{\mathfrak{R}^d} q(t, x; q_{i+1}^n, y) v_{i+1}^n(dy). \quad (2.22)$$

Here $v_{i+1}^n(dy)$ is a σ -finite measure on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ determined by $\rho(q_i^n, dx)$ and $\rho(q_{i+1}^n, dx)$. More precisely, on $[q_i^n, q_{i+1}^n]$, we consider Schrödinger's functional equation (see Definition 2.3) with end point distributions $\rho(q_i^n, dx)$ and $\rho(q_{i+1}^n, dx)$. $v_{i+1}^n(dy)$ is one of the solutions which is equivalent to $\rho(q_{i+1}^n, dx)$ (cf. (2.11) and Theorem 2.4).

The following lemma can be proved by Fleming's logarithmic transformation and reduces the argument on one-dimensional marginal distributions to that on the path space.

Lemma 2.5. Suppose that (A.0)–(A.2) hold. Then the infimum in (2.20) can be attained by the system of probability measures on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ which are one-dimensional marginal distributions of the \mathfrak{R}^d -valued Markovian Bernstein process ($P(\phi)_1$ -process)

$\{X^n(t)\}_{0 \leq t \leq T}$ on $(C([0, T]; \mathfrak{R}^d), B(C([0, T]; \mathfrak{R}^d)))$ which satisfies

$$P(X^n(q_i^n) \in dx) = \rho(q_i^n, dx) \quad \text{for all } i = 0, \dots, n, \quad (2.23)$$

$$dX^n(t) = b(t, X^n(t))^n dt + a(t, X^n(t))^{1/2} dW^n(t) \quad (2.24)$$

(in a weak sense) for b^n in (2.21), where $W^n(t)$ is a d -dimensional $\sigma[X^n(s); s \leq t]$ -Wiener process (cf. Theorem 2.4). We also have the following;

$$\begin{aligned} L_n &= \int_{C([0, T]; \mathfrak{R}^d)} d\mu^{X^n} \log(d\mu^{X^n}/d\mu^{X^0}) \\ &= \sum_{i=0}^{n-1} E[\log h_i^n(q_{i+1}^n, X^n(q_{i+1}^n))/h_i^n(q_i^n, X^n(q_i^n))], \end{aligned} \quad (2.25)$$

which is bounded with respect to $n \geq 1$. Here we denote by μ^{X^n} and μ^{X^0} the distributions of X^n and X^0 on $(C([0, T]; \mathfrak{R}^d), B(C([0, T]; \mathfrak{R}^d)))$, respectively.

Remark 2.3. We are considering the following change of measures:

$$\begin{aligned} &(d\mu^{X^n}/d\mu^{X^0})(X^0) \\ &= \{h_0^n(q_1^n, X^0(q_1^n))/h_0^n(0, X^0(0))\} \cdots \{h_{n-1}^n(T, X^0(T))/h_{n-1}^n(q_{n-1}^n, X^0(q_{n-1}^n))\}. \end{aligned} \quad (2.26)$$

This is true from (2.21) and the following; for $i = 0, \dots, n-1$,

$$\partial h_i^n(t, x)/\partial t = - \left(\sum_{j,k=1}^d a^{jk}(t, x) \partial^2 h_i^n(t, x) / \partial x_j \partial x_k \right) / 2 \quad \text{on } (q_i^n, q_{i+1}^n) \times \mathfrak{R}^d \quad (2.27)$$

from Lemma 2.3. This is the way to tie down X^0 so that (2.23) is satisfied.

Remark 2.4. Finiteness of L_n is equivalent to the following condition;

$$\sum_{i=0}^{n-1} E[|\log h_i^n(q_{i+1}^n, X^n(q_{i+1}^n))| + |\log h_i^n(q_i^n, X^n(q_i^n))|] < \infty \quad (2.28)$$

(cf. Csiszár [7]).

Let us prove Lemma 2.5.

Proof of Lemma 2.5. We prove Lemma 2.5 only when $n = 1$, since for $n \geq 2$, we can prove it, piecewise, in the same way as in the case in which $n = 1$.

Take $\phi \in C_c^\infty(|x| < 1; [0, \infty))$. For $R > 0$ and $\alpha > 0$, put

$$\phi_\alpha(x) \equiv \phi(x/\alpha) / \left(\int_{\mathfrak{R}^d} \phi(y/\alpha) dy \right), \quad (2.29)$$

$$h(t, x)^{R, \alpha}(t, x) \equiv \int_{\mathfrak{R}^d} q(t, x; T, y) \left(\int_{\mathfrak{R}^d} \phi_\alpha(y-z) [\min(v_T(z), R)] dz \right) dy, \quad (2.30)$$

$$h(t, x)^R \equiv \int_{\mathfrak{R}^d} q(t, x; T, y) [\min(v_T(y), R)] dy, \quad (2.31)$$

$$h(t, x) \equiv \int_{\mathfrak{R}^d} q(t, x; T, y) v_T(y) dy, \quad (2.32)$$

where we put

$$v_T(y) \equiv v_1^1(dy)/dy, \quad (2.33)$$

which exists from (A.0) (see (2.10) and (2.22)). Then

$$h^{R,\alpha} \in \bigcup_{0 < \varepsilon < T} C_b^{1,2}([0, T - \varepsilon] \times \mathfrak{R}^d; \mathfrak{R}), \quad (2.34)$$

$$h^{R,\alpha} \in C([0, T] \times \mathfrak{R}^d; \mathfrak{R}), \quad (2.35)$$

$$h(T, \cdot)^{R,\alpha} \in C_b^\infty(\mathfrak{R}^d; \mathfrak{R}) \quad (2.36)$$

(cf. Friedman [16 or 17]).

For any $\varepsilon > 0$, by Fleming's logarithmic transformation, we have the following dynamic programming equation; on $(0, T) \times \mathfrak{R}^d$,

$$\begin{aligned} & \partial \log(h(t, x)^{R,\alpha} + \varepsilon) / \partial t \\ &= - \left(\sum_{i,j=1}^d a^{ij}(t, x) \partial^2 \log(h(t, x)^{R,\alpha} + \varepsilon) / \partial x_i \partial x_j \right) / 2 \\ & \quad + \min_{u \in \mathfrak{R}^d} [\langle -u, \nabla_x \log(h(t, x)^{R,\alpha} + \varepsilon) \rangle + \langle a(t, x)^{-1} u, u \rangle / 2], \end{aligned} \quad (2.37)$$

since

$$\partial h(t, x)^{R,\alpha} / \partial t = - \left(\sum_{i,j=1}^d a^{ij}(t, x) \partial^2 h(t, x)^{R,\alpha} / \partial x_i \partial x_j \right) / 2. \quad (2.38)$$

Hence for any $0 < t < T$ and any $r \in \mathbf{A}_1$ (see Definition 2.4 for notations; in particular, b^r has nothing to do with b^n),

$$\begin{aligned} & \int_0^t \int_{\mathfrak{R}^d} \langle a(s, x)^{-1} b(s, x)^r, b(s, x)^r \rangle r(s, dx) ds / 2 \\ & \geq \int_0^t \int_{\mathfrak{R}^d} [\partial \log(h(s, x)^{R,\alpha} + \varepsilon) / \partial s \\ & \quad + \left(\sum_{i,j=1}^d a^{ij}(s, x) \partial^2 \log(h(s, x)^{R,\alpha} + \varepsilon) / \partial x_i \partial x_j \right) / 2 \\ & \quad + \langle b(s, x)^r, \nabla_x \log(h(s, x)^{R,\alpha} + \varepsilon) \rangle] r(s, dx) ds \\ & \quad \cdot \left(= \int_0^t \int_{\mathfrak{R}^d} [- \langle a(s, x) \nabla_x \log(h(s, x)^{R,\alpha} + \varepsilon), \nabla_x \log(h(s, x)^{R,\alpha} + \varepsilon) \rangle / 2 \right. \\ & \quad \left. + \langle b(s, x)^r, \nabla_x \log(h(s, x)^{R,\alpha} + \varepsilon) \rangle] r(s, dx) ds \right) \\ & = \int_{\mathfrak{R}^d} \log(h(t, x)^{R,\alpha} + \varepsilon) r(t, dx) - \int_{\mathfrak{R}^d} \log(h(0, x)^{R,\alpha} + \varepsilon) r(0, dx) \end{aligned} \quad (2.39)$$

from (0.1), (2.34), (2.37) and (2.38).

Let $t \rightarrow T$ in (2.39). Then

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(s, x)^{-1} b(s, x)^r, b(s, x)^r \rangle r(s, dx) ds / 2 \\ & \geq \int_{\mathfrak{R}^d} \log(h(T, x)^{R,\alpha} + \varepsilon) r(T, dx) - \int_{\mathfrak{R}^d} \log(h(0, x)^{R,\alpha} + \varepsilon) r(0, dx) \end{aligned} \quad (2.40)$$

from the weak continuity of $r(\cdot, dx)$ which implies the tightness of $r(\cdot, dx)$, and from (2.35)–(2.36).

Let $\alpha \rightarrow 0$ in (2.40). Then from the bounded convergence theorem and the constraints on $r(T, dx)$ and $r(0, dx)$,

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(s, x)^{-1} b(s, x)^r, b(s, x)^r \rangle r(s, dx) ds / 2 \\ & \geq E[\log \{ (h(T, X^1(T))^R + \varepsilon) / (h(0, X^1(0))^R + \varepsilon) \}] \\ & = E[\{ h(T, X^0(T)) / h(0, X^0(0)) \} \log \{ (h(T, X^0(T))^R + \varepsilon) / (h(0, X^0(0))^R + \varepsilon) \}] \end{aligned} \quad (2.41)$$

(cf. (2.26)). Let $\varepsilon \rightarrow 0$ and $R \rightarrow \infty$ in (2.41) at the same time. Then

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(s, x)^{-1} b(s, x)^r, b(s, x)^r \rangle r(s, dx) ds / 2 \\ & \geq E[\{ h(T, X^0(T)) / h(0, X^0(0)) \} \log \{ h(T, X^0(T)) / h(0, X^0(0)) \}] \\ & = E[\log \{ h(T, X^1(T)) / h(0, X^1(0)) \}] \end{aligned} \quad (2.42)$$

from Fatou's lemma. Here we used the following inequalities; if $h(T, X^0(T)) \geq R$,

$$\{ h(T, X^0(T)) / h(0, X^0(0)) \} \log \{ (h(T, X^0(T))^R + \varepsilon) / (h(0, X^0(0))^R + \varepsilon) \} \geq 0. \quad (2.43)$$

To get (2.43), from (2.31),

$$h(0, X^0(0))^R \leq R; \quad (2.44)$$

if $h(T, X^0(T)) \leq R$,

$$\begin{aligned} & \{ h(T, X^0(T)) / h(0, X^0(0)) \} \log \{ (h(T, X^0(T))^R + \varepsilon) / (h(0, X^0(0))^R + \varepsilon) \} \\ & \geq \{ h(T, X^0(T)) / h(0, X^0(0)) \} \log \{ (h(T, X^0(T)) + \varepsilon) / (h(0, X^0(0)) + \varepsilon) \}, \end{aligned} \quad (2.45)$$

since from (2.31) and (2.32),

$$h(0, X^0(0))^R \leq h(0, X^0(0)); \quad (2.46)$$

$$\begin{aligned} & \{ h(T, X^0(T)) / h(0, X^0(0)) \} \log \{ (h(T, X^0(T)) + \varepsilon) / (h(0, X^0(0)) + \varepsilon) \} \\ & \geq \begin{cases} 0 & \text{if } h(T, X^0(T)) \geq h(0, X^0(0)) \text{ or } h(T, X^0(T)) = 0; \\ -e^{-1} & \text{if } 0 < h(T, X^0(T)) \leq h(0, X^0(0)). \end{cases} \end{aligned} \quad (2.47)$$

Here we used the following inequality;

$$(x/y) \log \{ (x + \varepsilon) / (y + \varepsilon) \} \geq (x/y) \log(x/y) \geq -e^{-1} \quad \text{for all } y \geq x > 0. \quad (2.48)$$

Q.E.D.

3. Proof of Main Result

Let us prove Theorem 1.1. All the assumptions can be found in Sect. 1. First we outline the idea of the proof of Theorem 1.1. Consider $\rho(\cdot, dx)$ as a probability measure (on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$) valued function on $[0, T]$. For each $n \geq 1$, we approximate this function by a ‘‘piecewise linear’’ function. By a ‘‘piecewise linear’’ function, we

mean a probability measure (on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$) valued function on $[0, T]$ which attains the infimum in (2.20). As a minimizing function, we can take a system of one-dimensional marginal distributions of a Markovian Bernstein process $(P(\phi)_1\text{-process}) X^n(\cdot)$ (cf. Lemma 2.5) which converges, as $n \rightarrow +\infty$, to a Markov process in Theorem 1.1. To prove this convergence, we use Föllmer's entropy argument which inspired us Theorem 1.1 (cf. (Step 1) below). L_n in (2.20) is less than the infimum in (1.2) and converge to the infimum in (1.2) as $n \rightarrow \infty$ (cf. (Step 3) below). (1.5) can be proved by Föllmer's approach (cf. Föllmer [14]).

Proof of Theorem 3. Put

$$\begin{cases} \Omega \equiv C([0, T]; \mathfrak{R}^d) \\ L_n = \int_{\Omega} d\mu^{X^n} \log(d\mu^{X^n}/d\mu^{X^0}) \equiv H(\mu^{X^n}; \mu^{X^0}) \end{cases} \quad (3.1)$$

(cf. Lemma 2.5).

We divide the proof into 6 steps.

(Step 1). From Lemma 2.5, there exist a continuous semimartingale $X(\cdot)$ on $(C([0, T]; \mathfrak{R}^d), B(C([0, T]; \mathfrak{R}^d)))$ and a $\sigma[X(s); s \leq t]$ -adapted process $u(t, \omega)$ such that

$$dX(t) = u(t, \omega)dt + a(t, X(t))^{1/2}dW(t), \quad (3.2)$$

where $W(t)$ is a d -dimensional $\sigma[X(s); s \leq t]$ -Wiener process and that

$$H(\mu^X; \mu^{X^0}) = \int_0^T E[\langle a(t, X(t))^{-1}u(t, \omega), u(t, \omega) \rangle] dt/2 < \infty. \quad (3.3)$$

Let us prove (3.2)–(3.3). The following equality is true;

$$L_m = H(\mu^{X^m}; \mu^{X^0}) = H(\mu^{X^m}; \mu^{X^n}) + L_n \quad \text{for all } m \geq n, \quad (3.4)$$

since

$$\begin{aligned} L_m &= \int_{\Omega} d\mu^{X^m} \log(d\mu^{X^m}/d\mu^{X^0}) \\ &= \int_{\Omega} d\mu^{X^m} \log(d\mu^{X^m}/d\mu^{X^n}) + \int_{\Omega} d\mu^{X^m} \log(d\mu^{X^n}/d\mu^{X^0}), \end{aligned} \quad (3.5)$$

$$\begin{aligned} \int_{\Omega} d\mu^{X^m} \log(d\mu^{X^n}/d\mu^{X^0}) &= \sum_{i=0}^{n-1} E[\log h_i^n(q_{i+1}^n, X^m(q_{i+1}^n)) - \log h_i^n(q_i^n, X^m(q_i^n))] \\ &= \sum_{i=0}^{n-1} E[\log h_i^n(q_{i+1}^n, X^n(q_{i+1}^n)) - \log h_i^n(q_i^n, X^n(q_i^n))] \\ &= \int_{\Omega} d\mu^{X^n} \log(d\mu^{X^n}/d\mu^{X^0}), \end{aligned} \quad (3.6)$$

for all $m \geq n$. Here we used Remarks 2.3 and 2.4, and the following fact;

$$P(X^m(q_i^n) \in dx) = P(X^n(q_i^n) \in dx) = \rho(q_i^n, dx) \quad \text{for all } i = 0, \dots, n \quad (3.7)$$

for $m \geq n$ from (2.16).

Since $\{L_n\}_{n \geq 1}$ is a bounded monotone sequence (cf. Remark 2.2), L_n converges

as $n \rightarrow \infty$ and from (3.4),

$$\|\mu^{X^n} - \mu^{X^m}\|_V \leq \{2H(\mu^{X^m}; \mu^{X^n})\}^{1/2} \rightarrow 0 \quad \text{as } m \geq n \rightarrow \infty \quad (3.8)$$

(cf. Csizár [6] or Föllmer [14], p. 133, (3.4)) where $\|\cdot\|_V$ denotes a total variation norm.

Therefore there exists a probability measure μ^X on $(C([0, T]; \mathfrak{R}^d), B(C([0, T]; \mathfrak{R}^d)))$ such that

$$\|\mu^{X^n} - \mu^X\| \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (3.9)$$

$$H(\mu^X; \mu^{X^0}) < \infty, \quad (3.10)$$

since

$$\lim_{n \rightarrow \infty} H(\mu^{X^n}; \mu^{X^0}) \geq H(\mu^X; \mu^{X^0}), \quad (3.11)$$

from Fatou's lemma. (3.10) implies that μ^X is absolutely continuous with respect to μ^{X^0} and that μ^X is a probability law of a semimartingale $X(\cdot)$ which satisfies (3.2)–(3.3) for some $\sigma[X(s); s \leq t]$ -adapted $u(t, \omega)$ (cf. Liptser and Shiryaev [24], Theorem 7.11). Until now, we used Föllmer's entropy approach (cf. Föllmer [14], Sect. 2.1.4).

(Step 2). By the continuity of $\{X(t)\}_{0 \leq t \leq T}$ and (0.1), (1.4) holds. In fact, from the construction,

$$P(X(q) \in dx) = \rho(q, dx) \quad \text{for all } q \in Q_T \cup \{0, T\}. \quad (3.12)$$

For any $f \in C_b^2(\mathfrak{R}^d; \mathfrak{R})$ and any $0 \leq t \leq T$, there exists a sequence $\{q_l\}_{l \geq 1} \in Q_T$ which converges to t as $l \rightarrow \infty$ and

$$E[f(X(t))] = \lim_{l \rightarrow \infty} E[f(X(q_l))] = \lim_{l \rightarrow \infty} \int_{\mathfrak{R}^d} f(x) \rho(q_l, dx) = \int_{\mathfrak{R}^d} f(x) \rho(t, dx), \quad (3.13)$$

since from (0.1), $\int_{\mathfrak{R}^d} f(x) \rho(t, dx)$ is continuous.

(Step 3). We also have

$$E[u|(s, X(s))] = u(s, \omega). \quad (3.14)$$

In fact, since $P(X(t) \in dx) = \rho(t, dx) \in \mathbf{A}_n$ with

$$b(s, x)^\rho \equiv E[u|(s, X(s) = x)] \quad (3.15)$$

(cf. Sect. 0, (0.11)–(0.13)),

$$\int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^\rho, b(t, x)^\rho \rangle \rho(t, dx) dt / 2 \geq \lim_{n \rightarrow \infty} L_n \quad (3.16)$$

from (2.20). On the other hand,

$$\begin{aligned} \lim_{n \rightarrow \infty} L_n &\geq \int_0^T E[\langle a(t, X(t))^{-1} u(t, \omega), u(t, \omega) \rangle] dt / 2 \\ &\geq \int_0^T E[\langle a(t, X(t))^{-1} E[u|(t, X(t))], E[u|(t, X(t))] \rangle] dt / 2 \\ &= \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^\rho, b(t, x)^\rho \rangle \rho(t, dx) dt / 2, \end{aligned} \quad (3.17)$$

by (3.11) and Jensen's inequality (cf. Rockafellar [32]). From (3.16) and (3.17),

$$\begin{aligned} & \int_0^T E[\langle a(t, X(t))^{-1}u(t, \omega), u(t, \omega) \rangle] dt/2 \\ &= \int_0^T E[\langle a(t, X(t))^{-1}E[u|(t, X(t))], E[u|(t, X(t))] \rangle] dt/2, \end{aligned} \quad (3.18)$$

which implies (3.14) from the strict convexity of

$$\langle a(t, x)^{-1}u, u \rangle \quad \text{as a function of } u \quad (3.19)$$

for all $0 \leq t \leq T$ and $x \in \mathfrak{R}^d$ (from (A.2)).

(Step 4). Let us prove the existence and the uniqueness of the solution of the variational problem associated with $\{\rho(t, dx)\}_{0 \leq t \leq T}$. From (2.20) and (3.15)–(3.17), $b(t, x)^o$ is a solution $b(t, x)^o$ of the variational problem associated with $\{\rho(t, dx)\}_{0 \leq t \leq T}$. The uniqueness can be proved in the following way; for $b(t, x)' \in \mathbf{A}$ and $\alpha \in \mathfrak{R}$,

$$\alpha b(t, x)' + (1 - \alpha)b(t, x)^o \in \mathbf{A} \quad (3.20)$$

(see Definition 1.2) and

$$\begin{aligned} & d \left(\int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(\alpha b(t, x)' + (1 - \alpha)b(t, x)^o), \right. \\ & \left. \alpha b(t, x)' + (1 - \alpha)b(t, x)^o \rangle \rho(t, dx) dt \right) / d\alpha|_{\alpha=0} = 0, \end{aligned} \quad (3.21)$$

this is,

$$\int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x)^o, b(t, x)' - b(t, x)^o \rangle \rho(t, dx) dt = 0. \quad (3.22)$$

Hence

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x)^o, b(t, x)^o \rangle \rho(t, dx) dt/2 \\ &= - \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x)^o, b(t, x)^o \rangle \rho(t, dx) dt/2 \\ &+ \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x)^o, b(t, x)' \rangle \rho(t, dx) dt \\ &= - \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(b(t, x)^o - b(t, x)'), b(t, x)^o - b(t, x)' \rangle \rho(t, dx) dt/2 \\ &+ \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x)', b(t, x)' \rangle \rho(t, dx) dt/2, \end{aligned} \quad (3.23)$$

which implies the uniqueness of b^o .

(Step 5). The uniqueness and the Markov property of $X(\cdot)$ can be derived from (3.2), (3.10) and (3.14). In fact, (3.10) implies the existence and the uniqueness of

Maruyama–Girsanov density $d\mu^X/d\mu^{X^0}$ (cf. Liptser and Shiryaev [24], Theorem 7.10) and therefore the Markov property of X from that of X^0 , (3.2) and (3.14).

(Step 6). (1.5) can be proved in the same way as in Föllmer [13].

Put, for $0 \leq t \leq T$,

$$\begin{cases} \hat{X}(t) \equiv X(T-t), \\ \hat{X}^0(t) \equiv X^0(T-t). \end{cases} \quad (3.24)$$

Then from Haussmann and Pardoux's results [19], $\hat{X}^0(t)$ is a weak solution of the following stochastic differential equation (cf. Stroock and Varadhan [37]);

$$\begin{cases} d\hat{X}^0(t) = \hat{b}(t, \hat{X}^0(t))dt + \hat{a}(t, \hat{X}^0(t))^{1/2}d\hat{W}^0(t) & \text{on } [0, T], \\ P(\hat{X}^0(0) \in dx) = P(X^0(T) \in dx), \end{cases} \quad (3.25)$$

where $\hat{W}^0(t)$ is a d -dimensional $\sigma[\hat{X}^0(s); 0 \leq s \leq t]$ -Wiener process. Here we put

$$\hat{a}(t, x) \equiv a(T-t, x), \quad (3.26)$$

$$\hat{b}(t, x) \equiv p(T-t, x)^{-1} \left(\sum_{j=1}^d \partial(\hat{a}^{ij}(t, x)p(T-t, x))/\partial x_j \right)_{i=1}^d, \quad (3.27)$$

$$p(t, x) \equiv \int_{\mathfrak{R}^d} q(0, y; t, x)\rho(0, dy) \quad (3.28)$$

(see below (2.5)).

The remaining part of the proof is the same as Föllmer [13]. Q.E.D.

Remark 3.1.

$$H(\mu^X; \mu^{X^n}) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (3.29)$$

that is,

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(b(t, x)^o - b(t, x)^n), b(t, x)^o - b(t, x)^n \rangle \rho(t, dx) dt \\ &= \sum_{i=0}^{n-1} \int_{q_i^n}^{q_{i+1}^n} \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(b(t, x)^o - a(t, x)\nabla_x \log h_i^n(t, x)), \\ & \quad \cdot b(t, x)^o - a(t, x)\nabla_x \log h_i^n(t, x) \rangle \rho(t, dx) dt \\ & \rightarrow 0 \quad \text{as } n \rightarrow \infty, \end{aligned} \quad (3.30)$$

which implies that b^o is not always equal to b (cf. Proposition 3.1). This is true, since

$$\begin{aligned} & H(\mu^X; \mu^{X^n}) \\ &= \int_{\Omega} d\mu^X \log(d\mu^X/d\mu^{X^n}) \\ &= \int_{\Omega} d\mu^X \log(d\mu^X/d\mu^{X^0}) - \int_{\Omega} d\mu^X \log(d\mu^{X^n}/d\mu^{X^0}) \\ &= H(\mu^X; \mu^{X^0}) - H(\mu^{X^n}; \mu^{X^0}) \\ & \rightarrow 0 \quad \text{as } n \rightarrow \infty \end{aligned} \quad (3.31)$$

from (3.16) and (3.17). Here we used the same argument as in (3.6).

A mean forward derivative of a variational process can be approximated as in (3.30) and is not always unique. But when is b for which (0.2) holds equal to b^0 ? We conclude this section with the following proposition.

Proposition 3.1. *Suppose that (A.0)–(A.2) hold. Then b for which (0.2) holds is equal to b^0 in Theorem 1.1 if there exists a sequence of functions $\{f^n(t, x)\}$ for which $f^n(t, x) \in C_b^{1,2}([q_i^n, q_{i+1}^n] \times \mathfrak{R}^d; \mathfrak{R}) \cap C_b([q_i^n, q_{i+1}^n] \times \mathfrak{R}^d; \mathfrak{R})$ for all $i = 0, \dots, n-1$ and $f^n(q_i^n, x) \in C_b^2(\mathfrak{R}^d; \mathfrak{R})$ for all $i = 1, \dots, n$ such that*

$$\lim_{n \rightarrow \infty} \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(b(t, x) - a(t, x)\nabla_x f^n(t, x)), b(t, x) - a(t, x)\nabla_x f^n(t, x) \rangle \rho(t, dx) dt = 0. \quad (3.32)$$

Moreover, if for any compact subset $K \subset [0, T] \times \mathfrak{R}^d$, there exists a constant $C_K > 0$ such that

$$\rho(t, x) \geq C_K \quad \text{on } K, \quad (3.33)$$

then there exists a function $\Phi(\cdot, \cdot)$ for which $\Phi(t, \cdot) \in H_{\text{loc}}^1(\mathfrak{R}^d)$ for almost all $0 \leq t \leq T$ such that

$$a^{-1}(t, x)b(t, x) = \nabla_x \Phi(t, x) \quad dt \, dx\text{-a.e.} \quad (3.34)$$

Here $H_{\text{loc}}^1(\mathfrak{R}^d)$ denotes the completion of $C_0^\infty(\mathfrak{R}^d; \mathfrak{R})$ with respect to the seminorms; for $\varphi \in C_0^\infty(\mathfrak{R}^d; \mathfrak{R})$

$$\int_{|x| \leq N} (\varphi(x)^2 + |\nabla \varphi(x)|^2) dx \quad \text{for all } N \geq 1. \quad (3.35)$$

Proof. Let us prove the first part.

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(b(t, x)^0 - a(t, x)\nabla_x f^n(t, x)), b(t, x)^0 - a(t, x)\nabla_x f^n(t, x) \rangle \rho(t, dx) dt \\ &= \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x)^0, b(t, x)^0 \rangle \rho(t, dx) dt \\ & \quad - 2 \int_0^T \int_{\mathfrak{R}^d} \langle b(t, x)^0, \nabla_x f^n(t, x) \rangle \rho(t, dx) dt \\ & \quad + \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)\nabla_x f^n(t, x), \nabla_x f^n(t, x) \rangle \rho(t, dx) dt \\ & \leq \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}b(t, x), b(t, x) \rangle \rho(t, dx) dt \\ & \quad - 2 \int_0^T \int_{\mathfrak{R}^d} \langle b(t, x), \nabla_x f^n(t, x) \rangle \rho(t, dx) dt \\ & \quad + \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)\nabla_x f^n(t, x), \nabla_x f^n(t, x) \rangle \rho(t, dx) dt \\ &= \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1}(b(t, x) - a(t, x)\nabla_x f^n(t, x)), b(t, x) - a(t, x)\nabla_x f^n(t, x) \rangle \rho(t, dx) dt \\ & \rightarrow 0 \quad \text{as } n \rightarrow \infty \end{aligned} \quad (3.36)$$

from (3.32). Here we used the optimality of b^o and the fact that

$$\int_0^T \int_{\mathfrak{R}^d} \langle b(t, x)^o - b(t, x), \nabla_x f^n(t, x) \rangle \rho(t, dx) dt = 0, \quad (3.37)$$

since $b^o \in \mathbf{A}$ (see the argument below (2.40)).

Let us prove (3.34). From (3.33),

$$a(t, x)^{-1} b(t, x) = \lim_{n \rightarrow \infty} \nabla_x f^n(t, x) \quad \text{in } L_{\text{loc}}^2([0, T] \times \mathfrak{R}^d), \quad (3.38)$$

where $L_{\text{loc}}^2([0, T] \times \mathfrak{R}^d)$ denotes the set of functions $f(\cdot, \cdot): [0, T] \times \mathfrak{R}^d \mapsto \mathfrak{R}$ such that for any compact subset $F \subset [0, T] \times \mathfrak{R}^d$,

$$\iint_F f(t, x)^2 dt dx < +\infty. \quad (3.39)$$

Therefore for almost all $0 \leq t \leq T$, there exists a generalized function $G(t, \cdot)$ such that

$$a(t, x)^{-1} b(t, x) = \nabla_x G(t, x) \quad dx dt\text{-a.e.} \quad (3.40)$$

From (A.1), $G(t, x)$ can be identified with some function $\Phi(t, x)$ for which $\Phi(t, \cdot) \in H_{\text{loc}}^1(\mathfrak{R}^d)$ for almost all $0 \leq t \leq T$. (cf. Maz'ja [26], p. 23, Corollary). Q.E.D.

4. Schrödinger's Problem: Revisited

In this section we consider the application of Theorem 1.1. In this section, we fix $T > 0$. All the assumptions can be found in Sect. 1.

Let us reformulate Schrödinger's Problem as we mentioned in Remark 1.3.

Let \mathfrak{p} be the set of systems of Borel probability measures $\{r(t, dx)\}_{0 \leq t \leq T}$ which satisfies the following; for any $f \in C_b^{1,2}([0, T] \times \mathfrak{R}^d; \mathfrak{R})$ and any $0 \leq s \leq t \leq T$,

$$\begin{aligned} & \int_{\mathfrak{R}^d} f(t, x) r(t, dx) - \int_{\mathfrak{R}^d} f(s, x) r(s, dx) \\ &= \int_s^t \int_{\mathfrak{R}^d} \left[\partial f(u, x) / \partial u + \left(\sum_{i,j=1}^d a^{ij}(u, x) \partial^2 f(u, x) / \partial x_i \partial x_j \right) / 2 \right. \\ & \quad \left. + \sum_{i=1}^d b^i(u, x) \partial f(u, x) / \partial x_i \right] r(u, dx) du, \end{aligned} \quad (4.1)$$

for some $b(t, x)^r \equiv (b^i(t, x))^r_{i=1}$, where $a(t, x) \equiv (a^{ij}(t, x))^r_{i,j=1}$ satisfies (A.2).

Put, for each $r \in \mathfrak{p}$ and b^r ,

$$E(r, b^r) \equiv \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^r, b(t, x)^r \rangle r(t, dx) dt. \quad (4.2)$$

Notice that there may be many b^r for each $r \in \mathfrak{p}$ and in the next theorem the infimum is taken over all $r \in \mathfrak{p}$ and all b^r .

Theorem 4.1. *Suppose that (A.2) holds. Then for any probability measures $\rho_0(dx)$ and $\rho_T(dx) \equiv \rho_T(x) dx$ on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ for which*

$$\inf \{ E(r, b^r); r(0, dx) = \rho_0(dx), r(T, dx) = \rho_T(dx), r \in \mathfrak{p}, b^r \} < \infty, \quad (4.3)$$

the infimum in (4.3) can be attained by the system of one-dimensional marginal distributions of an \mathfrak{R}^d -valued Markov process $\{X(t) \equiv X(t, \omega)\}_{0 \leq t \leq T}$ on $(C([0, T]; \mathfrak{R}^d), B(C([0, T]; \mathfrak{R}^d)))$ which satisfies

$$dX(t) = a(t, X(t))\nabla_x \log h(t, X(t))dt + a(t, X(t))^{1/2}dW(t) \quad (4.4)$$

in a weak sense (cf. Stroock and Varadhan [37]) (see (2.11)), where $W(t)$ is a d -dimensional $\sigma[X(s); s \leq t]$ -Wiener process, and

$$\begin{cases} P(X(0) \in dx) = \rho_0(dx), \\ P(X(T) \in dx) = \rho_T(dx), \end{cases} \quad (4.5)$$

$$\begin{aligned} & \inf \{E(r, b^r); r(0, dx) = \rho_0(dx), r(T, dx) = \rho_T(dx), r \in \mathfrak{p}, b^r\} \\ & = \int_0^T E[\langle a(t, X(t))^{-1} \{a(t, X(t))\nabla_x \log h(t, X(t))\}, \\ & \quad a(t, X(t))\nabla_x \log h(t, X(t)) \rangle] dt. \end{aligned} \quad (4.6)$$

Moreover, if (A.0) is satisfied for all $r \in \mathfrak{p}$ which attains the infimum in (4.3), then the infimum in (4.3) can be attained only by the system of one-dimensional marginal distributions of $X(\cdot)$ which satisfies (4.4)–(4.5).

Proof. We only have to prove the uniqueness from Lemma 2.5, where our case is that in which $n = 1$. Notice that the set $\{r \in \mathfrak{p}; r(0, dx) = \rho_0(dx), r(T, dx) = \rho_T(dx)\}$ is a set \mathbf{A}_1 in Definition 2.4 with $\rho(q_0^1, dx) = \rho_0(dx)$ and $\rho(q_1^1, dx) = \rho_T(dx)$.

From the proof of the existence (2.39)–(2.48) (especially (2.39)), for any $r \in \mathbf{A}_1$ which attains the infimum in L_1 (see Definition 2.4),

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0, R \rightarrow \infty} \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} (b(t, x)^r - a(t, x)\nabla_x \log(h(t, x)^R + \varepsilon)), \\ & \quad b(t, x)^r - a(t, x)\nabla_x \log(h(t, x)^R + \varepsilon) \rangle r(t, dx) dt = 0. \end{aligned} \quad (4.7)$$

Therefore

$$b(t, x)^r = a(t, x)\nabla_x \log h(t, x) (\equiv b(t, x)^1) \quad r(t, dx) dt \text{ a.e.} \quad (4.8)$$

(cf. (2.21)) from (A.2), which implies the uniqueness of the mean forward derivative. Notice that r in $b(t, x)^r$ is not a number but a probability measure in \mathbf{A}_1 .

Consider the variational problem associated with $\{r(t, dx)\}_{0 \leq t \leq T}$. Then from Theorem 1.1, there exists a unique optimal mean forward derivative $b^{r, o}$ such that

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^r, b(t, x)^r \rangle r(t, dx) dt \\ & \quad \geq \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^{r, o}, b(t, x)^{r, o} \rangle r(t, dx) dt. \end{aligned} \quad (4.9)$$

But from the definition of b^r ,

$$\begin{aligned} & \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^{r, o}, b(t, x)^{r, o} \rangle r(t, dx) dt \\ & \quad \geq \int_0^T \int_{\mathfrak{R}^d} \langle a(t, x)^{-1} b(t, x)^r, b(t, x)^r \rangle r(t, dx) dt. \end{aligned} \quad (4.10)$$

From (4.9)–(4.10) and the uniqueness of the optimal mean forward derivative,

$$b(t, x)^r = b(t, x)^{r \circ}.$$

Moreover, again from Theorem 1.1, $\{r(t, dx)\}_{0 \leq t \leq T}$ is a system of the one-dimensional marginal distributions of the Markov diffusion process $\{X^r(t)\}_{0 \leq t \leq T}$ which satisfies the following stochastic differential equation (in a weak sense) (cf. Stroock and Varadhan [37]);

$$dX^r(t) = a(t, X^r(t))\nabla_x \log h(t, X^r(t))dt + a(t, X^r(t))^{1/2}dW^r(t),$$

with

$$H(\mu^{X^r}; \mu^{X^0}) < \infty.$$

(4.13) implies the absolute continuity of μ^{X^r} with respect to μ^{X^0} and the uniqueness of the weak solution of (4.12) (cf. Liptser and Shiryaev [24], Theorem 7.10) and hence that of $\{r(t, dx)\}_{0 \leq t \leq T}$. Q.E.D.

Definition 4.1. *The problem to find a Markov process which satisfies (4.5)–(4.6) from given two Borel probability measures $\rho_0(dx)$ and $\rho_T(dx)$ on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ is called Schrödinger's Problem.*

Remark 4.1. As we mentioned in Remark 1.3, our class of systems of one-dimensional marginal distributions under consideration is those of Borel probability measures on $(\mathfrak{R}^d, B(\mathfrak{R}^d))$ and larger than that considered by Zambrini [38] (cf. also Dai pra [8], Föllmer [14], and Schrödinger [34, 35]). They only considered semimartingales. The uniqueness in Theorem 4.1 is equivalent to that of a solution $\{\rho(t, dx)\}_{0 \leq t \leq T}$ of the following weak forward equation; for any $f \in C_b^{1,2}([0, T] \times \mathfrak{R}^d, \mathfrak{R})$ and any $0 \leq s \leq t \leq T$,

$$\begin{aligned} & \int_{\mathfrak{R}^d} f(t, x)\rho(t, dx) - \int_{\mathfrak{R}^d} f(s, x)\rho(s, dx) \\ &= \int_s^t \int_{\mathfrak{R}^d} \left[\partial f(u, x)/\partial u + \left(\sum_{i,j=1}^d a^{ij}(u, x)\partial^2 f(u, x)/\partial x_i \partial x_j \right) \right] / 2 \\ & \quad + \langle a(u, x)\nabla_x \log h(u, x), \nabla_x f(u, x) \rangle \rho(u, dx) du. \end{aligned} \quad (4.14)$$

If we restrict the class of systems of one-dimensional marginal distributions under consideration to those of semimartingales, then (4.14) has a unique solution in that class (see (4.12)–(4.13) and Remark 1.1). In fact, to prove the existence, we do not need to use the condition (A.0) (cf. the proof of Lemma 2.5).

Remark 4.2. The Markov process in Theorem 4.1 is a Markovian Bernstein process ($P(\phi)_1$ -process) constructed from $q(s, x; t, y)$ and end point distributions $\rho_0(dx), \rho_T(dx)$ (see Definition 2.3–Theorem 2.2).

Acknowledgement. This is a part of the author's dissertation under the supervision of Professor Wendell H. Fleming at Brown University. The author wishes to thank him for the useful discussions and the constant encouragement. The author would also like to thank the referee for pointing out the relation between Bernstein processes and local Markov random fields.

References

1. Bernstein, S.: Sur les liaisons entre les grandeurs aléatoires. *Verh. des Intern. Mathematikerkongr. Zürich. Band I* (1932)
2. Beurling, A.: An automorphism of product measures. *Ann. Math.* **72**, 189–200 (1960)
3. Carlen, E. A.: Conservative diffusions. *Commun. Math. Phys.* **94**, 293–315 (1984)
4. Carlen, E. A.: Existence and sample path properties of the diffusions in Nelson's stochastic mechanics. In: Alberverio, S., Blanchard, Ph., Streit, L. (eds.) *Proc. of the BiBoS-Symp., "Stochastic Processes—Mathematics and Physics," Lecture Notes in Math., Vol. 1158*, pp. 25–51. Berlin, Heidelberg, New York: Springer 1986
5. Carmona, R.: Probabilistic construction of Nelson processes. *Taniguchi Symp. PMMP Katata*. 55–81 (1985)
6. Csiszár, I.: Information-Type measure of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **2**, 299–318 (1967)
7. Csiszár, I.: I-Divergence geometry of probability distributions and minimization problems. *Ann. Probab.* **3**, 146–158 (1975)
8. Dai Pra, P.: Stochastic control approach to reciprocal diffusion processes. To appear in *Appl. Math. Optim.* (1989)
9. Doob, J. L.: A Markov chain theorem. *Probability and Statistics (The Harald Cramér Volume)*. New York: Wiley 1959
10. Dynkin, E. B.: *Markov Processes I*. Berlin, Heidelberg, New York: Springer 1965
11. Dynkin, E. B.: *Markov Processes II*. Berlin, Heidelberg, New York: Springer 1965
12. Fleming, W. H., Rishel, D. E.: *Deterministic and stochastic optimal control*. Berlin, Heidelberg, New York: Springer 1975
13. Föllmer, H.: Time reversal on Wiener space. In: Alberverio, S., Blanchard, Ph., Streit, L. (eds.) *Proc. of the BiBoS-Symp., "Stochastic Processes—Mathematics and Physics." Lecture Notes in Math., Vol. 1158*, pp. 119–129. Berlin, Heidelberg, New York: Springer 1986
14. Föllmer, H.: Random fields and diffusion processes. In: Hennequin, P. L. (ed.) *École d'Été de Probabilités Saint-Flour XV–XVII, 1985–87. Lecture Notes in Math., Vol. 1362*, pp. 101–203. Berlin, Heidelberg, New York: Springer 1988
15. Fortet, R.: Résolution d'un Système d'équations de M. Schroedinger. *J. Math. Pures Appl.* **IX**, 83–105 (1940)
16. Friedman, A.: *Partial differential equations of parabolic type*. New Jersey: Prentice-Hall Englewood Cliffs 1964
17. Friedman, A.: *Stochastic differential equations and applications. Vol. I*. New York: Academic Press 1975
18. Glimm, J., Jaffe, A.: *Quantum physics. A functional integral point of view*. Berlin, Heidelberg, New York: Springer 1987
19. Haussmann, U. G., Pardoux, E.: Time Reversal of Diffusions. *Ann. Probab.* **14**, 1188–1205 (1986)
20. Ikeda, N., Watanabe, S.: *Stochastic differential equations and diffusion processes*. New York: North-Holland 1981
21. Jamison, B.: Reciprocal processes. *Z. Wahrsch. Verw. Gebiete* **30**, 65–86 (1974)
22. Jamison, B.: The Markov processes of Schrödinger. *Z. Wahrsch. Verw. Gebiete* **32**, 323–331 (1975)
23. Krylov, N. V.: On quasi-diffusional processes. *Theory Probab. Appl.* **11**, 373–389 (1966)
24. Liptser, R. S., Shiryaev, A. N.: *Statistics of random processes in*. Berlin, Heidelberg, New York: Springer 1977
25. Loève, M.: *Probability theory*. Princeton: Van Nostrand 1963
26. Maz'ja, V. G.: *Sobolev spaces*. Berlin, Heidelberg, New York: Springer 1985
27. Meyer, P. A., Zheng, W. A.: Tightness criteria for laws of semimartingales. *Ann. Inst. Henri Poincaré* **20**, 353–372 (1984)
28. Mikami, T.: *Small random perturbations of dynamical systems and Markov processes from marginal distributions*. Thesis, Brown University 1989
29. Nagasawa, M.: Transformations of diffusion and Schrödinger processes. *Probab. Th. Rel. Fields* **82**, 109–136 (1989)
30. Nelson, E.: *Dynamical theories of Brownian motion*. Princeton, NJ: Princeton University Press 1967

31. Nelson, E.: Quantum fluctuations. Princeton, NJ: Princeton University Press 1984
32. Rockafellar, R. T.: Convex analysis. Princeton mathematical series, Vol. 28. Princeton, NJ: Princeton University Press 1970
33. Sakurai, J. J.: Modern quantum mechanics. California: Benjamin/Cummings 1985
34. Schrödinger, E.: Über die Umkehrung der Naturgesetze. Sitz. Ber. der Preuss. Akad. Wissen. Berlin Phys. Math. **144**, 144–153 (1931)
35. Schrödinger, E.: Théorie relativiste de l'électron et l'interprétation de la mécanique quantique. Ann. Inst. Henri Poincaré **2**, 269–310 (1932)
36. Simon, B.: Functional integration and quantum physics. New York: Academic Press 1979
37. Stroock, D. W., Varadhan, S. R. S.: Multidimensional diffusion processes. Berlin, Heidelberg, New York: Springer 1979
38. Zambrini, J. C.: Variational Processes. In: Alberverio, S., Casati, G., Merlini, D. (eds.). Proc. of the 1st Ascona-Como International Conference, "Stochastic Processes in Classical and Quantum Systems," Lecture Notes in Physics, Vol. **262**, pp. 517–529. Berlin, Heidelberg, New York: Springer 1986
39. Zheng, W. A.: Tightness results for laws of diffusion processes application to stochastic mechanics. Ann. Inst. Henri Poincaré **21**, 103–124 (1985)

Communicated by A. Jaffe