Elect. Comm. in Probab. **7** (2002) 199–213

ELECTRONIC COMMUNICATIONS in PROBABILITY

OPTIMAL CONTROL FOR ABSOLUTELY CONTINUOUS STOCHASTIC PROCESSES AND THE MASS TRANSPORTATION PROBLEM

TOSHIO MIKAMI

Department of Mathematics, Hokkaido University Sapporo 060-0810, Japan

email: mikami@math.sci.hokudai.ac.jp

submitted January 15, 2002 *Final version accepted* October 28, 2002

AMS 2000 Subject classification: Primary: 93E20

Absolutely continuous stochastic process, mass transportation problem, Salisbury's problem, Markov control, zero-noise limit

Abstract

We study the optimal control problem for $\mathbf{R}^d$-valued absolutely continuous stochastic processes with given marginal distributions at every time. When $d = 1$, we show the existence and the uniqueness of a minimizer which is a function of a time and an initial point. When $d > 1$, we show that a minimizer exists and that minimizers satisfy the same ordinary differential equation.

1 Introduction

Monge-Kantorovich problem (MKP for short) plays a crucial role in many fields and has been studied by many authors (see [2, 3, 7, 10, 12, 20] and the references therein).

Let $h : \mathbf{R}^d \mapsto [0, \infty)$ be convex, and Q_0 and Q_1 be Borel probability measures on $\mathbf{R}^d$, and put

$$\mu_h(Q_0, Q_1) := \inf E[\int_0^1 h\left(\frac{d\phi(t)}{dt}\right)dt], \tag{1.1}$$

where the infimum is taken over all absolutely continuous stochastic processes $\{\phi(t)\}_{0\le t\le 1}$ for which $P\phi(t)^{-1} = Q_t (t = 0, 1)$. (In this paper we use the same notation P for different probability measures, for the sake of simplicity, when it is not confusing.)

As a special case of MKPs, we introduce the following problem (see e.g. [2, 3] and also [18]).

Does there exist a minimizer $\{\phi^o(t)\}_{0\le t\le 1}$, of (1.1), which is a function of t and $\phi^o(0)$?

Suppose that there exist $p \in L^1([0, 1]\times\mathbf{R}^d : \mathbf{R}, dtdx)$ and $b(t, x) \in L^1([0, 1]\times\mathbf{R}^d : \mathbf{R}^d, p(t, x)dtdx)$ such that the following holds: for any $f \in C_o^\infty(\mathbf{R}^d)$ and any $t \in [0, 1]$,

$$\int_{\mathbf{R}^d} f(x)(p(t, x) - p(0, x))dx = \int_0^t ds \int_{\mathbf{R}^d} < \nabla f(x), b(s, x) > p(s, x)dx,$$

$$p(t,x) \geq 0 \quad dx-\text{a.e.}, \qquad \int_{\mathbf{R}^d} p(t,x)dx = 1. \tag{1.2}$$

Here $< \cdot, \cdot >$ denotes the inner product in $\mathbf{R}^d$ and $\nabla f(x) := (\partial f(x)/\partial x_i)_{i=1}^d$.
Put, for $n \geq 1$,

$$\mathbf{e}_n := \inf\{E[\int_0^1 h\left(\frac{dY(t)}{dt}\right)dt] : \{Y(t)\}_{0\leq t\leq 1} \in A_n\}, \tag{1.3}$$

where A_n is the set of all absolutely continuous stochastic processes $\{Y(t)\}_{0\leq t\leq 1}$ for which $P(Y(t) \in dx) = p(t,x)dx$ for all $t = 0, 1/n, \cdots, 1$.
Then the minimizer of $\mathbf{e}_n$ can be constructed by those of

$$\mu_{\frac{h(n\cdot)}{n}}\left(p\left(\frac{k}{n},x\right)dx, p\left(\frac{k+1}{n},x\right)dx\right) \quad (k = 0, \cdots, n-1)$$

(see (1.1) for notation). As $n \to \infty$, $\mathbf{e}_n$ formally converges to

$$\mathbf{e} := \inf\{E[\int_0^1 h\left(\frac{dY(t)}{dt}\right)dt] : \{Y(t)\}_{0\leq t\leq 1} \in A\}, \tag{1.4}$$

where A is the set of all absolutely continuous stochastic processes $\{Y(t)\}_{0\leq t\leq 1}$ for which $P(Y(t) \in dx) = p(t,x)dx$ for all $t \in [0,1]$.
In this sense, the minimizer of $\mathbf{e}$ can be considered as the continuum limit of those of $\mathbf{e}_n$ as $n \to \infty$.
In this paper, instead of $h(u)$, we would like to consider more general function $L(t,x;u) : [0,1] \times \mathbf{R}^d \times \mathbf{R}^d \mapsto [0,\infty)$ which is convex in u, and study the minimizers of

$$\mathbf{e}^0 := \inf\{E[\int_0^1 L\left(t, Y(t); \frac{dY(t)}{dt}\right)dt] : \{Y(t)\}_{0\leq t\leq 1} \in A\}. \tag{1.5}$$

Remark 1 *It is easy to see that the set A_n is not empty, but it is not trivial to show that the set A is not empty if b in (1.2) is not smooth. As a similar problem, that of the construction of a Markov diffusion process $\{X(t)\}_{0\leq t\leq 1}$ such that $PX(t)^{-1}$ satisfies a given Fokker-Planck equation with nonsmooth coefficients is known and has been studied by many authors (see [4], [5], [15], [19] and the references therein).*

We would also like to point out that (1.1) and (1.5) can be formally considered as the zero-noise limits of h-path processes and variational processes, respectively, when $h = L = |u|^2$ (see [8] and [15], respectively).
More generally, we have the following.
Let $(\Omega, \mathbf{B}, P)$ be a probability space, and $\{\mathbf{B}_t\}_{t\geq 0}$ be a right continuous, increasing family of sub σ-fields of $\mathbf{B}$, and X_o be a $\mathbf{R}^d$-valued, $\mathbf{B}_0$-adapted random variable such that $PX_o^{-1}(dx) = p(0,x)dx$, and $\{W(t)\}_{t\geq 0}$ denote a d-dimensional $(\mathbf{B}_t)$-Wiener process (see e.g. [11] or [13]).
For $\varepsilon > 0$ and a $\mathbf{R}^d$-valued $(\mathbf{B}_t)$-progressively measurable $\{u(t)\}_{0\leq t\leq 1}$, put

$$X^{\varepsilon,u}(t) := X_o + \int_0^t u(s)ds + \varepsilon W(t) \quad (t \in [0,1]). \tag{1.6}$$

Put also

$$\mathbf{e}^{\varepsilon} := \inf\{E[\int_0^1 L(t, X^{\varepsilon,u}(t); u(t))dt] : \{u(t)\}_{0\le t\le 1} \in A^{\varepsilon}\} \quad (\varepsilon > 0), \tag{1.7}$$

where $A^{\varepsilon} := \{\{u(t)\}_{0\le t\le 1} : P(X^{\varepsilon,u}(t) \in dx) = p(t,x)dx(0 \le t \le 1)\}$; and

$$\tilde{\mathbf{e}}^{\varepsilon} := \inf\{\int_0^1 \int_{\mathbf{R}^d} L(t, y; B(t, y))p(t, y)dtdy : B \in \tilde{A}^{\varepsilon}\} \quad (\varepsilon \ge 0), \tag{1.8}$$

where $\tilde{A}^{\varepsilon}$ is the set of all $B(t, x) : [0, 1] \times \mathbf{R}^d \mapsto \mathbf{R}^d$ for which the following holds: for any $f \in C_o^{\infty}(\mathbf{R}^d)$ and any $t \in [0, 1]$,

$$\begin{aligned} &\int_{\mathbf{R}^d} f(x)(p(t,x) - p(0,x))dx \\ = &\int_0^t ds \int_{\mathbf{R}^d} \Big(\frac{\varepsilon^2 \triangle f(x)}{2} + < \nabla f(x), B(s,x) > \Big) p(s,x)dx. \end{aligned}$$

Then we expect that the following holds:

$$\mathbf{e}^{\varepsilon} = \tilde{\mathbf{e}}^{\varepsilon} \to \mathbf{e}^0 = \tilde{\mathbf{e}}^0 \qquad (\text{as } \varepsilon \to 0). \tag{1.9}$$

In this paper we show that the set A is not empty and (1.9) holds, and that a minimizer of $\mathbf{e}^0$ exists when the cost function $L(t, x; u)$ grows at least of order of $|u|^2$ as $u \to \infty$ (see Theorem 1 in section 2).
We also show that the minimizers satisfy the same ordinary differential equation (ODE for short) when L is strictly convex in u (see Theorem 2 in section 2). (In this paper we say that a function $\{\psi(t)\}_{0\le t\le 1}$ satisfies an ODE if and only if it is absolutely continuous and $d\psi(t)/dt$ is a function of t and $\psi(t)$, dt-a.e..)
When $d = 1$, we show the uniqueness of the minimizer of $\mathbf{e}^0$ (see Corollary 1 in section 2).
Since a stochastic process which satisfies an ODE is not always nonrandom, we would also like to know if the minimizer is a function of a time and an initial point. In fact, the following is known as Salisbury's problem (SP for short).

Is a continuous strong Markov process which is of bounded variation in time a function of an initial point and a time?

If $x(t)_{0\le t\le 1}$ is a $\mathbf{R}$-valued strong Markov process, and if there exists a Borel measurable function f, on $\mathbf{R}$, such that $x(t) = x(0) + \int_0^t f(x(s))ds$ $(0 \le t \le 1)$, then SP has been solved positively by Çinlar and Jacod (see [6]). When $d > 1$, a counter example is known (see [21]).
When $d = 1$, we give a positive answer to SP for time-inhomogeneous stochastic processes (see Proposition 2 in section 4). This is a slight generalization of [6] where they made use of the result on time changes of Markov processes, in that the stochastic processes under consideration are time-inhomogeneous and need not be Markovian. In particular, we show, when $d = 1$, that $\{Y(t)\}_{0\le t\le 1}, \in A$, which satisfies an ODE is unique and nonrandom. It will be used to show that the unique minimizer of $\mathbf{e}^0$ is a function of an initial point and of a time when $d = 1$ (see Corollary 1 and Theorem 3 in section 2).

Remark 2 *When $d > 1$, $\{Y(t)\}_{0\le t\le 1}, \in A$, which satisfies an ODE is not unique (see Proposition 1 in section 2).*

When $L(t,x;u)=|u|^2$ and $p(t,x)$ satisfies the Fokker-Planck equation with sufficiently smooth coefficients, the optimization problem (1.5) was considered in [16] where the minimizer exists uniquely and is a function of a time and an initial point, and where we used a different approach which depends on the form of $L(t,x;u)=|u|^2$.
Our main tool in the proof is the weak convergence method, the result on the construction of a Markov diffusion process from a family of marginal distributions, and the theory of Copulas. In section 2 we state our main result. We first consider the case where a cost function $L(t,x;u)$ grows at least of order of $|u|^2$ as $u \to \infty$ and $d \geq 1$. Next we restrict our attention to the case where L is a function of u and $d=1$. The proof is given in section 3. We discuss SP in section 4.

2 Main result

In this section we give our main result.
We state assumptions before we state the result when $d \geq 1$.
(H.0). $\tilde{\mathbf{e}}^0$ is finite (see (1.8) for notation).
(H.1). $L(t,x;u):[0,1]\times\mathbf{R}^d\times\mathbf{R}^d \mapsto [0,\infty)$ is convex in u, and as $h,\delta \downarrow 0$,

$$R(h,\delta):=\sup\left\{\frac{L(t,x;u)-L(s,y;u)}{1+L(s,y;u)}:|t-s|<h,|x-y|<\delta,u\in\mathbf{R}^d\right\}\downarrow 0.$$

(H.2). There exists $q \geq 2$ such that the following holds:

$$0<\liminf_{|u|\to\infty}\frac{\inf\{L(t,x;u):(t,x)\in[0,1]\times\mathbf{R}^d\}}{|u|^q}, \tag{2.10}$$

$$\sup\left\{\frac{\sup_{z\in\partial_u L(t,x;u)}|z|}{(1+|u|)^{q-1}}:(t,x,u)\in[0,1]\times\mathbf{R}^d\times\mathbf{R}^d\right\}\equiv C_{\nabla L}<\infty, \tag{2.11}$$

where $\partial_u L(t,x;u):=\{z\in\mathbf{R}^d : L(t,x;v)-L(t,x;u)\geq <z,v-u> \text{ for all } v\in\mathbf{R}^d\}$ ($t\in[0,1]$, x, $u\in\mathbf{R}^d$).
(H.3). $p(t,\cdot)$ is absolutely continuous dt-a.e., and for q in (H.2),

$$\int_0^1\int_{\mathbf{R}^d}\left|\frac{\nabla_x p(t,x)}{p(t,x)}\right|^q p(t,x)dtdx<\infty. \tag{2.12}$$

Remark 3 *If (H.0) does not hold, then $\mathbf{e}^0$ in (1.5) is infinite. (H.1) implies the continuity of $L(\cdot,\cdot;u)$ for each $u\in\mathbf{R}^d$. (H.2) holds if $L(t,x;u)=|u|^q$. We need (H.3) to make use of the result on the construction of a Markov diffusion process of which the marginal distribution at time t is $p(t,x)dx$ ($0\leq t\leq 1$). (2.3) holds if $b(t,x)$ in (1.2) is twice continuously differentiable with bounded derivatives up to the second order, and if $p(0,x)$ is absolutely continuous, and if the following holds:*

$$\int_{\mathbf{R}^d}\left|\frac{\nabla_x p(0,x)}{p(0,x)}\right|^q p(0,x)dx<\infty. \tag{2.13}$$

The following theorem implies the existence of a minimizer of $\mathbf{e}^0$ (see (1.5)-(1.8) for notations).

Theorem 1 *Suppose that (H.0)-(H.3) hold. Then the sets A^ε $(\varepsilon > 0)$ and A are not empty, and the following holds:*

$$\tilde{\mathbf{e}}^\varepsilon = \mathbf{e}^\varepsilon \to \mathbf{e}^0 = \tilde{\mathbf{e}}^0 \quad (\text{as } \varepsilon \to 0). \tag{2.14}$$

In particular, for any $\{u^\varepsilon(t)\}_{0\le t\le 1}, \in A^\varepsilon(\varepsilon > 0)$, for which

$$\lim_{\varepsilon\to 0} E[\int_0^1 L(t, X^{\varepsilon,u^\varepsilon}(t); u^\varepsilon(t))dt] = \mathbf{e}^0, \tag{2.15}$$

$\{\{X^{\varepsilon,u^\varepsilon}(t)\}_{0\le t\le 1}\}_{\varepsilon>0}$ is tight in $C([0,1] : \mathbf{R}^d)$, and any weak limit point of $\{X^{\varepsilon,u^\varepsilon}(t)\}_{0\le t\le 1}$ as $\varepsilon \to 0$ is a minimizer of $\mathbf{e}^0$.

The following theorem implies the uniqueness of the minimizer of $\tilde{\mathbf{e}}^0$ and that of the ODE which is satisfied by the minimizers of $\mathbf{e}^0$.

Theorem 2 *Suppose that (H.0)-(H.3) hold. Then for any minimizer $\{X(t)\}_{0\le t\le 1}$ of $\mathbf{e}^0$, $b^X(t,x) := E[dX(t)/dt|(t, X(t) = x)]$ is a minimizer of $\tilde{\mathbf{e}}^0$. Suppose in addition that L is strictly convex in u. Then $\tilde{\mathbf{e}}^0$ has the unique minimizer $b^o(t,x)$ and the following holds: for any minimizer $\{X(t)\}_{0\le t\le 1}$ of $\mathbf{e}^0$,*

$$X(t) = X(0) + \int_0^t b^o(s, X(s))ds \quad \text{for all } t \in [0,1], \quad a.s.. \tag{2.16}$$

Remark 4 *By Theorems 1 and 2, if (H.0) with $L = |u|^2$ and (H.3) with $q = 2$ hold, then there exists a stochastic process $\{X(t)\}_{0\le t\le 1}, \in A$, which satisfies an ODE.*

Since $b \in \tilde{A}^0$ is not always the gradient, in x, of a function, the following implies that the set $\tilde{A}^0$ does not always consist of only one point.

Proposition 1 *Suppose that $L = |u|^2$, and that (H.0) and (H.3) with $q = 2$ hold, and that for any $M > 0$,*

$$ess.inf\{p(t,x) : t \in [0,1], |x| \le M\} > 0. \tag{2.17}$$

Then the unique minimizer of $\tilde{\mathbf{e}}^0$ can be written as $\nabla_x V(t,x)$, where $V(t,\cdot) \in H^1_{loc}(\mathbf{R}^d : \mathbf{R})$ dt-a.e..

We next consider the one-dimensional case. Put

$$F_t(x) := \int_{(-\infty,x]} p(t,y)dy \quad (t \in [0,1], x \in \mathbf{R}),$$

$$F_t^{-1}(u) := \sup\{y \in \mathbf{R} : F_t(y) < u\} \quad (t \in [0,1], 0 < u < 1).$$

(H.3)'. $d = 1$, and $F_t(x)$ is differentiable and has the locally bounded first partial derivatives on $[0,1] \times \mathbf{R}$.
By Proposition 2 in section 4, we obtain the following.

Corollary 1 *Suppose that (H.0)-(H.3) and (H.3)' hold, and that L is strictly convex in u. Then the minimizer $\{X(t)\}_{0\le t\le 1}$ of $\mathbf{e}^0$ is unique. Moreover, $\lim_{s\in\mathbf{Q}\cap[0,1],s\to t} F_s^{-1}(F_0(X(0)))$ exists and is equal to $X(t)$ for all $t \in [0,1]$ a.s..*

The theory of copulas allows us to treat a different set of assumptions by a different method (see (1.3)-(1.4) for notations).
(H.0)'. $\{\mathbf{e}_n\}_{n\geq 1}$ is bounded.
(H.1)'. $h : \mathbf{R} \mapsto [0,\infty)$ is even and convex.
(H.2)'. There exists $r > 1$ such that the following holds:

$$0 < \liminf_{|u|\to\infty} \frac{h(u)}{|u|^r}. \tag{2.18}$$

(H.3)". $d = 1$, and $p(t,x)$ is positive on $[0,1] \times \mathbf{R}$.

Theorem 3 *Suppose that (H.0)'-(H.2)' and (H.3)" hold. Then* $\{F_t^{-1}(F_0(x))\}_{0\leq t\leq 1}$ *on* $(\mathbf{R}, \mathbf{B}(\mathbf{R}), p(0,x)dx)$ *belongs to the set A and is a minimizer of* $\mathbf{e}$*. Suppose in addition that (H.3)' holds. Then* $\{F_t^{-1}(F_0(x))\}_{0\leq t\leq 1}$ *is the unique minimizer, of* $\mathbf{e}$*, that satisfies an ODE.*

Remark 5 *If* $\{\mathbf{e}_n\}_{n\geq 1}$ *is unbounded, then so is* $\mathbf{e}$*. By (H.1)',* $\{\overline{X}(t) := F_t^{-1}(F_0(x))\}_{0\leq t\leq 1}$ *satisfies the following (see e.g. [20, Chap. 3.1]): for any t and* $s \in [0,1]$,

$$\mu_h(p(s,x)dx, p(t,x)dx) = E_0[h(\overline{X}(t) - \overline{X}(s))] \tag{2.19}$$

(see (1.1) for notation), where we put $P_0(dx) := p(0,x)dx$*. Indeed,*

$$\overline{X}(t) = F_t^{-1}(F_s(\overline{X}(s))) \tag{2.20}$$

since for a distribution F on $\mathbf{R}$,

$$F(F^{-1}(u)) = u \quad (0 < u < 1) \tag{2.21}$$

(see e. g. [17]).

3 Proof of the result

In this section we prove the result given in section 2.
Before we give the proof of Theorem 1, we state and prove three technical lemmas.

Lemma 1 *Suppose that (H.2) holds. Then for any* $\varepsilon > 0$, $\tilde{\mathbf{e}}^\varepsilon = \mathbf{e}^\varepsilon$.

(Proof). For any $B^\varepsilon \in \tilde{A}^\varepsilon$ for which $\int_0^1 \int_{\mathbf{R}^d} L(t,x;B^\varepsilon(t,x))p(t,x)dtdx$ is finite, there exists a Markov process $\{Z^\varepsilon(t)\}_{0\leq t\leq 1}$ such that the following holds:

$$Z^\varepsilon(t) = X_o + \int_0^t B^\varepsilon(s, Z^\varepsilon(s))ds + \varepsilon W(t), \tag{3.22}$$

$$P(Z^\varepsilon(t) \in dx) = p(t,x)dx \quad (0 \leq t \leq 1), \tag{3.23}$$

since $\int_0^1 \int_{\mathbf{R}^d} |B^\varepsilon(t,x)|^2 p(t,x)dtdx$ is finite by (H.2) (see [4] and [5]). This implies that $\{B^\varepsilon(t, Z^\varepsilon(t))\}_{0\leq t\leq 1} \in A^\varepsilon$, and that the following holds:

$$\int_0^1 \int_{\mathbf{R}^d} L(t,x;B^\varepsilon(t,x))p(t,x)dtdx = \int_0^1 E[L(t,Z^\varepsilon(t);B^\varepsilon(t,Z^\varepsilon(t)))]dt, \tag{3.24}$$

from which $\mathbf{e}^{\varepsilon} \le \tilde{\mathbf{e}}^{\varepsilon}$.
We next show that $\mathbf{e}^{\varepsilon} \ge \tilde{\mathbf{e}}^{\varepsilon}$.
For any $\{u^{\varepsilon}(t)\}_{0\le t\le 1} \in A^{\varepsilon}$, $b^{\varepsilon,u^{\varepsilon}}(t,x) := E[u^{\varepsilon}(t)|(t, X^{\varepsilon,u^{\varepsilon}}(t)=x)] \in \tilde{A}^{\varepsilon}$.
Indeed, for any $f \in C_o^{\infty}(\mathbf{R}^d)$ and any $t\in[0,1]$, by the Itô formula,

$$\begin{aligned} & \int_{\mathbf{R}^d} f(x)(p(t,x)-p(0,x))dx = E[f(X^{\varepsilon,u^{\varepsilon}}(t)) - f(X^{\varepsilon,u^{\varepsilon}}(0))] \qquad (3.25)\\ = & \int_0^t E\Big[\frac{\varepsilon^2}{2}\triangle f(X^{\varepsilon,u^{\varepsilon}}(s)) + < \nabla f(X^{\varepsilon,u^{\varepsilon}}(s)), u^{\varepsilon}(s) > \Big]ds \\ = & \int_0^t E\Big[\frac{\varepsilon^2}{2}\triangle f(X^{\varepsilon,u^{\varepsilon}}(s)) + < \nabla f(X^{\varepsilon,u^{\varepsilon}}(s)), b^{\varepsilon,u^{\varepsilon}}(s, X^{\varepsilon,u^{\varepsilon}}(s)) > \Big]ds \\ = & \int_0^t ds \int_{\mathbf{R}^d}\Big(\frac{\varepsilon^2}{2}\triangle f(x) + < \nabla f(x), b^{\varepsilon,u^{\varepsilon}}(s,x) > \Big)p(s,x)dx. \end{aligned}$$

The following completes the proof: by Jensen's inequality,

$$\begin{aligned} & \int_0^1 E[L(t, X^{\varepsilon,u^{\varepsilon}}(t); u^{\varepsilon}(t))]dt \qquad (3.26)\\ \ge & \int_0^1 E[L(t, X^{\varepsilon,u^{\varepsilon}}(t); b^{\varepsilon,u^{\varepsilon}}(t, X^{\varepsilon,u^{\varepsilon}}(t)))]dt \\ = & \int_0^1 \int_{\mathbf{R}^d} L(t,x; b^{\varepsilon,u^{\varepsilon}}(t,x))p(t,x)dtdx. \end{aligned}$$

Q. E. D.

The following lemma can be shown by the standard argument and the proof is omitted (see [13, p. 17, Theorem 4.2 and p. 33, Theorem 6.10]).

Lemma 2 *For any* $\{u^{\varepsilon}(t)\}_{0\le t\le 1} \in A^{\varepsilon}$ $(\varepsilon > 0)$ *for which* $\{E[\int_0^1 |u^{\varepsilon}(t)|^2 dt]\}_{\varepsilon>0}$ *is bounded,* $\{\{X^{\varepsilon,u^{\varepsilon}}(t)\}_{0\le t\le 1}\}_{\varepsilon>0}$ *is tight in* $C([0,1]:\mathbf{R}^d)$.

Lemma 3 *For any* $\{u^{\varepsilon_n}(t)\}_{0\le t\le 1} \in A^{\varepsilon_n}$ $(n\ge 1)$ $(\varepsilon_n \to 0$ *as* $n\to\infty)$ *such that* $\{E[\int_0^1 |u^{\varepsilon_n}(t)|^2 dt]\}_{n\ge 1}$ *is bounded and that* $\{X_n(t) := X^{\varepsilon_n,u^{\varepsilon_n}}(t)\}_{0\le t\le 1}$ *weakly converges as* $n\to\infty$*, the weak limit* $\{X(t)\}_{0\le t\le 1}$ *in* $C([0,1]:\mathbf{R}^d)$ *is absolutely continuous.*

(Proof). We only have to show the following: for any $\delta > 0$ and any $m\ge 2$, $n \ge 1$ and any $s_{i,j}, t_{i,j} \in \mathbf{Q}$ for which $0 \le s_{i,j} \le t_{i,j} \le s_{i,j+1} \le t_{i,j+1} \le 1$ $(1\le i\le n, 1\le j\le m-1)$ and for which $\sum_{j=1}^m |t_{i,j}-s_{i,j}| \le \delta$ $(1\le i\le n)$,

$$E[\max_{1\le i\le n}(\sum_{j=1}^m |X(t_{i,j}) - X(s_{i,j})|)^2] \le \delta \liminf_{k\to\infty} E[\int_0^1 |u^{\varepsilon_k}(t)|^2 dt]. \qquad (3.27)$$

Indeed, by the monotone convergence theorem and by the continuity of $\{X(t)\}_{0\le t\le 1}$, (3.6) implies that, for all $m\ge 2$,

$$E[\sup\{(\sum_{j=1}^{m}|X(t_j)-X(s_j)|)^2 : \sum_{j=1}^{m}|t_j-s_j|\le\delta \\ ,0\le s_j\le t_j\le s_{j+1}\le t_{j+1}\le 1(1\le j\le m-1)\}] \\ \le \quad \delta\liminf_{k\to\infty}E[\int_0^1|u^{\varepsilon_k}(t)|^2dt]. \tag{3.28}$$

The left hand side of (3.7) converges, as $m\to\infty$, to

$$E[\sup\{(\sum_{j=1}^{m}|X(t_j)-X(s_j)|)^2 : \sum_{j=1}^{m}|t_j-s_j|\le\delta, m\ge 2 \\ ,0\le s_j\le t_j\le s_{j+1}\le t_{j+1}\le 1(1\le j\le m-1)\}] \tag{3.29}$$

since the integrand on the left hand side of (3.7) is nondecreasing in m.
Hence by Fatou's lemma,

$$\lim_{\delta\to 0}(\sup\{(\sum_{j=1}^{m}|X(t_j)-X(s_j)|)^2 : \sum_{j=1}^{m}|t_j-s_j|\le\delta, m\ge 2 \\ ,0\le s_j\le t_j\le s_{j+1}\le t_{j+1}\le 1(1\le j\le m-1)\})=0 \quad \text{a.s.}, \tag{3.30}$$

since the integrand in (3.8) is nondecreasing in $\delta>0$ and henceforth is convergent as $\delta\to 0$. To complete the proof, we prove (3.6). By Jensen's inequality, for $i=1,\cdots,n$ for which $\sum_{j=1}^{m}|t_{i,j}-s_{i,j}|>0$,

$$(\sum_{j=1}^{m}|X(t_{i,j})-X(s_{i,j})|)^2 \\ \le \quad (\sum_{j=1}^{m}|t_{i,j}-s_{i,j}|)\sum_{1\le j\le m, s_{i,j}<t_{i,j}}\left|\frac{X(t_{i,j})-X(s_{i,j})}{t_{i,j}-s_{i,j}}\right|^2(t_{i,j}-s_{i,j}). \tag{3.31}$$

Put $A_{mn}:=\{t_{i,j},s_{i,j};1\le i\le n,1\le j\le m\}$ and $\{t_k\}_{1\le k\le\#(A_{mn})}:=A_{mn}$ so that $t_k<t_{k+1}$ for $k=1,\cdots,\#(A_{mn})-1$, where $\#(A_{mn})$ denotes the cardinal number of the set A_{mn}. Then, by Jensen's inequality,

$$\sum_{1\le j\le m, s_{i,j}<t_{i,j}}\left|\frac{X(t_{i,j})-X(s_{i,j})}{t_{i,j}-s_{i,j}}\right|^2(t_{i,j}-s_{i,j}) \\ \le \quad \sum_{1\le k\le\#(A_{mn})-1}\left|\frac{X(t_k)-X(t_{k+1})}{t_{k+1}-t_k}\right|^2(t_{k+1}-t_k). \tag{3.32}$$

The following completes the proof: for any $k=1,\cdots,\#(A_{mn})-1$,

$$E\left[\left|\frac{X(t_k)-X(t_{k+1})}{t_{k+1}-t_k}\right|^2\right] \le \liminf_{\ell\to\infty} E\left[\left|\frac{X_\ell(t_k)-X_\ell(t_{k+1})}{t_{k+1}-t_k}\right|^2\right] \le \frac{1}{t_{k+1}-t_k}\liminf_{\ell\to\infty} E[\int_{t_k}^{t_{k+1}}|u^{\varepsilon_\ell}(t)|^2dt]. \tag{3.33}$$

Q. E. D.

We prove Theorem 1 by Lemmas 1-3.
(Proof of Theorem 1). The proof of (2.5) is divided into the following:

$$\limsup_{\varepsilon\to 0}\tilde{\mathbf{e}}^\varepsilon \le \tilde{\mathbf{e}}^0, \tag{3.34}$$

$$\liminf_{\varepsilon\to 0}\mathbf{e}^\varepsilon \ge \mathbf{e}^0, \tag{3.35}$$

since $\mathbf{e}^\varepsilon=\tilde{\mathbf{e}}^\varepsilon$ by Lemma 1, and since $\mathbf{e}^0\ge\tilde{\mathbf{e}}^0$ in the same way as in the proof of the inequality $\mathbf{e}^\varepsilon\ge\tilde{\mathbf{e}}^\varepsilon$ (see (3.4)-(3.5)).
We first prove (3.13). For $B\in\tilde{A}^0$ for which $\int_0^1\int_{\mathbf{R}^d}L(t,x;B(t,x))p(t,x)dtdx$ is finite and $\varepsilon>0$, $B(t,x)+\varepsilon^2\nabla p(t,x)/(2p(t,x))\in\tilde{A}^\varepsilon$.
Indeed, for any $f\in C_o^\infty(\mathbf{R}^d)$ and any $t\in[0,1]$,

$$\begin{aligned}&\int_{\mathbf{R}^d}f(x)(p(t,x)-p(0,x))dx\\ =\ &\int_0^t ds\int_{\mathbf{R}^d}<\nabla f(x),B(s,x)>p(s,x)dx\\ =\ &\int_0^t ds\int_{\mathbf{R}^d}\Bigl(\frac{\varepsilon^2}{2}\triangle f(x)+\Bigl\langle\nabla f(x),B(s,x)+\frac{\varepsilon^2\nabla p(s,x)}{2p(s,x)}\Bigr\rangle\Bigr)p(s,x)dx.\end{aligned}$$

For any $t\in[0,1]$, x, u, $v\in\mathbf{R}^d$, and $z\in\partial_u L(t,x;u+v)$, by (2.2),

$$\begin{aligned}&L(t,x;u+v)\le L(t,x;u)-<z,v>\\ \le\ &L(t,x;u)+C_{\nabla L}(1+|u+v|)^{q-1}|v|.\end{aligned}\tag{3.36}$$

Putting $u=B(t,x)$ and $v=\varepsilon^2\nabla p(t,x)/(2p(t,x))$ in (3.15), we have

$$\begin{aligned}\tilde{\mathbf{e}}^\varepsilon\ &\le \int_0^1\int_{\mathbf{R}^d}L\Bigl(t,x;B(t,x)+\frac{\varepsilon^2\nabla p(t,x)}{2p(t,x)}\Bigr)p(t,x)dtdx\\ &\le \int_0^1\int_{\mathbf{R}^d}C_{\nabla L}\Bigl(1+\Bigl|B(t,x)+\frac{\varepsilon^2\nabla p(t,x)}{2p(t,x)}\Bigr|\Bigr)^{q-1}\Bigl|\frac{\varepsilon^2\nabla p(t,x)}{2p(t,x)}\Bigr|p(t,x)dtdx\\ &\qquad+\int_0^1\int_{\mathbf{R}^d}L(t,x;B(t,x))p(t,x)dtdx\\ &\to \int_0^1\int_{\mathbf{R}^d}L(t,x;B(t,x))p(t,x)dtdx \quad (\text{as } \varepsilon\to 0)\end{aligned}\tag{3.37}$$

by (2.1) and (H.3), where we used the following in the last line of (3.16):

$$\frac{q-1}{q}+\frac{1}{q}=1.$$

Next we prove (3.14). By Lemmas 2-3, we only have to show the following: for any $\{u^{\varepsilon_n}(t)\}_{0\le t\le 1}\in A^{\varepsilon_n}$ $(n\ge 1)$ ($\varepsilon_n\to 0$ as $n\to\infty$) for which $\{X_n(t):=X^{\varepsilon_n,u^{\varepsilon_n}}(t)\}_{0\le t\le 1}$ weakly converges, as $n\to\infty$, to a stochastic process $\{X(t)\}_{0\le t\le 1}$, and for which $\{E[\int_0^1 L(t,X_n(t);u^{\varepsilon_n}(t))dt]\}_{n\ge 1}$ is bounded,

$$\liminf_{n\to\infty} E[\int_0^1 L(t,X_n(t);u^{\varepsilon_n}(t))dt]\ge E[\int_0^1 L\left(t,X(t);\frac{dX(t)}{dt}\right)dt]. \tag{3.38}$$

We prove (3.17). For $\alpha\in(0,1)$ and $\delta>0$,

$$\begin{aligned}
&E[\int_0^1 L(t,X_n(t);u^{\varepsilon_n}(t))dt] \qquad (3.39)\\
\ge\ &\frac{1}{1+R(\alpha,\delta)}E[\int_0^{1-\alpha} dsL\left(s,X_n(s);\frac{1}{\alpha}\int_s^{s+\alpha}u^{\varepsilon_n}(t)dt\right)\\
&;\sup_{0\le t,s\le 1,|t-s|<\alpha}|X_n(t)-X_n(s)|<\delta]-R(\alpha,\delta).
\end{aligned}$$

Indeed, if $\sup_{0\le t,s\le 1,|t-s|<\alpha}|X_n(t)-X_n(s)|<\delta$, then for $s\in[0,1-\alpha]$, by Jensen's inequality and (H.1),

$$\begin{aligned}
&L\left(s,X_n(s);\frac{1}{\alpha}\int_s^{s+\alpha}u^{\varepsilon_n}(t)dt\right)\le\frac{1}{\alpha}\int_s^{s+\alpha}L(s,X_n(s);u^{\varepsilon_n}(t))dt \qquad (3.40)\\
\le\ &R(\alpha,\delta)+\frac{1+R(\alpha,\delta)}{\alpha}\int_s^{s+\alpha}L(t,X_n(t);u^{\varepsilon_n}(t))dt.
\end{aligned}$$

Hence putting $u=\int_s^{s+\alpha}u^{\varepsilon_n}(t)dt/\alpha$ and $v=(X_n(s+\alpha)-X_n(s)-\int_s^{s+\alpha}u^{\varepsilon_n}(t)dt)/\alpha$ in (3.15), we have, from (3.18),

$$\begin{aligned}
&E[\int_0^1 L(t,X_n(t);u^{\varepsilon_n}(t))dt] \qquad (3.41)\\
\ge\ &\frac{1}{1+R(\alpha,\delta)}E[\int_0^{1-\alpha}L\left(s,X_n(s);\frac{X_n(s+\alpha)-X_n(s)}{\alpha}\right)ds\\
&;\sup_{0\le t,s\le 1,|t-s|<\alpha}|X_n(t)-X_n(s)|<\delta]\\
&-E[\int_0^{1-\alpha}C_{\nabla L}\left(1+\left|\frac{X_n(s+\alpha)-X_n(s)}{\alpha}\right|\right)^{q-1}\\
&\times\left|\frac{\varepsilon_n}{\alpha}(W(s+\alpha)-W(s))\right|ds]-R(\alpha,\delta).
\end{aligned}$$

Letting $n\to\infty$ and then $\alpha\to 0$ and $\delta\to 0$ in (3.20), we obtain (3.17).

Indeed, by Skorohod's theorem (see e.g. [13]), taking a new probability space, we can assume that $\{X_n(t)\}_{0\le t\le 1}$ converges, as $n\to\infty$, to $\{X(t)\}_{0\le t\le 1}$ in sup norm, a.s., and that the following holds: for any $\beta\in(0,\delta/3)$, by (H.1),

$$\begin{aligned}
&(1+R(0,\beta))E[\int_0^{1-\alpha} L\left(s,X_n(s);\frac{X_n(s+\alpha)-X_n(s)}{\alpha}\right)ds\\
&\qquad ;\sup_{0\le t,s\le 1,|t-s|<\alpha}|X_n(t)-X_n(s)|<\delta]\\
\ge\ & E[\int_0^{1-\alpha} L\left(s,X(s);\frac{X_n(s+\alpha)-X_n(s)}{\alpha}\right)ds;\sup_{0\le t\le 1}|X(t)-X_n(t)|<\beta\\
&\quad ,\sup_{0\le t,s\le 1,|t-s|<\alpha}|X(t)-X(s)|<\beta]-R(0,\beta).
\end{aligned}$$

The liminf of the right-hand side of this inequality as $n\to\infty$, and $\alpha\to 0$ and then $\beta\to 0$ is dominated by $E[\int_0^1 L(s,X(s);dX(s)/ds)ds]$ from below by Fatou's lemma. The second mean value on the right hand side of (3.20) can be shown to converge to zero as $n\to\infty$ in the same way as in (3.16) by (2.1).
(H.0) and (2.5) implies that the set A and A^ε $(\varepsilon>0)$ are not empty.
(2.5) and (3.17) completes the proof.

Q. E. D.

(Proof of Theorem 2). $b^X(t,x)$ is a minimizer of $\tilde{\mathbf{e}}^0$ by (2.5) in the same way as in (3.4)-(3.5). We prove the uniqueness of the minimizer of $\tilde{\mathbf{e}}^0$. Suppose that $b^o(t,x)$ is also a minimizer of $\tilde{\mathbf{e}}^0$. Then for any $\lambda\in(0,1)$, $\lambda b^X+(1-\lambda)b^o\in\tilde{A}^0$, and

$$\begin{aligned}
\tilde{\mathbf{e}}^0 &\le \int_0^1\int_{\mathbf{R}^d} L(t,y;\lambda b^X(t,y)+(1-\lambda)b^o(t,y))p(t,y)dtdy \\
&\le \lambda\int_0^1\int_{\mathbf{R}^d} L(t,y;b^X(t,y))p(t,y)dtdy\\
&\quad +(1-\lambda)\int_0^1\int_{\mathbf{R}^d} L(t,y;b^o(t,y))p(t,y)dtdy=\tilde{\mathbf{e}}^0.
\end{aligned}\tag{3.42}$$

By the strict convexity of L in u,

$$b^X(t,x)=b^o(t,x),\quad p(t,x)dtdx-a.e.. \tag{3.43}$$

We prove (2.7). Since L is strictly convex in u, the following holds:

$$\frac{dX(t)}{dt}=b^X(t,X(t))\quad dtdP-a.e. \tag{3.44}$$

by (2.5) (see (3.5)). By (3.22),

$$\begin{aligned}
&E[\sup_{0\le t\le 1}|X(t)-X(0)-\int_0^t b^o(s,X(s))ds|] \\
\le\ & \int_0^1 E[|b^X(s,X(s))-b^o(s,X(s))|]ds=0.
\end{aligned}\tag{3.45}$$

Q. E. D.

(Proof of Proposition 1). From [15], $\tilde{\mathbf{e}}^\varepsilon = \mathbf{e}^\varepsilon$ for $\varepsilon > 0$, and the minimizer of $\tilde{\mathbf{e}}^\varepsilon$ can be written as $\nabla_x \Phi^\varepsilon(t,x)$, where $\Phi^\varepsilon(t,\cdot) \in H^1_{loc}(\mathbf{R}^d : \mathbf{R})$ dt-a.e.. Since $\{\nabla_x \Phi^\varepsilon\}_{0<\varepsilon<1}$ is strongly bounded in $L^2([0,1]\times\mathbf{R}^d : \mathbf{R}^d, p(t,x)dtdx)$ by (2.5), it is weakly compact in $L^2([0,1]\times\mathbf{R}^d : \mathbf{R}^d, p(t,x)dtdx)$ (see [9, p. 639]). We denote a weak limit point by Ψ. Then Ψ is the unique minimizer of $\tilde{\mathbf{e}}^0$. Indeed, $\Psi \in \tilde{A}^0$, and by (2.5) and Fatou's lemma,

$$\tilde{\mathbf{e}}^0 = \lim_{\varepsilon\to 0}\int_0^1\int_{\mathbf{R}^d} |\nabla_x \Phi^\varepsilon(t,y)|^2 p(t,y)dtdy \tag{3.46}$$
$$\geq \int_0^1\int_{\mathbf{R}^d} |\Psi(t,y)|^2 p(t,y)dtdy \geq \tilde{\mathbf{e}}^0.$$

In particular, $\{\nabla_x \Phi^\varepsilon\}_{0<\varepsilon<1}$ converges, as $\varepsilon \to 0$, to Ψ, strongly in $L^2([0,1]\times\mathbf{R}^d : \mathbf{R}^d, p(t,x)dtdx)$, which completes the proof in the same way as in [15, Proposition 3.1].

Q. E. D.

Remark 6 *If $V(t,x)$ and $p(t,x)$ in Proposition 1 are sufficiently smooth, then*

$$\nabla_x \Phi^\varepsilon(t,x) = \nabla_x V(t,x) + \frac{\varepsilon^2 \nabla_x p(t,x)}{2p(t,x)}$$

(see [16, section 1]).

(Proof of Theorem 3). Put for $t \in [0,1]$, $x \in \mathbf{R}$ and $n \geq 1$,

$$Y(t,x) = F_t^{-1}(F_0(x)), \tag{3.47}$$
$$Y_n(t,x) = Y\left(\frac{[nt]}{n}, x\right) + n\left(t - \frac{[nt]}{n}\right)\left(Y\left(\frac{[nt]+1}{n}, x\right) - Y\left(\frac{[nt]}{n}, x\right)\right), \tag{3.48}$$

where $[nt]$ denotes the integer part of nt.
Then by (H.3)", $Y(\cdot,x) \in C([0,1] : \mathbf{R})$, $P_0(dx) := p(0,x)dx - a.s.$, and

$$\lim_{n\to\infty} Y_n(t,x) = Y(t,x) \quad (0 \leq t \leq 1), \quad P_0 - a.s., \tag{3.49}$$

and

$$\mathbf{e}_n = E_0[\int_0^1 h\left(\frac{dY_n(t,x)}{dt}\right)dt] \quad (n \geq 1) \tag{3.50}$$

(see Remark 5 in section 2 and [11, p. 35, Exam. 8.1]).
Hence in the same way as in the proof of Lemma 3, we can show that the following holds: for any $\delta > 0$

$$E_0[\sup\{(\sum_{j=1}^m |Y(t_j,x) - Y(s_j,x)|)^r : \sum_{j=1}^m |t_j - s_j| \leq \delta, m \geq 2 \tag{3.51}$$

$$
\begin{aligned}
&\quad ,0 \le s_j \le t_j \le s_{j+1} \le t_{j+1} \le 1 (1 \le j \le m-1)\}] \\
\le &\quad \delta^{r-1} \liminf_{n\to\infty} E_0[\int_0^1 \left|\frac{dY_n(t,x)}{dt}\right|^r dt],
\end{aligned}
$$

which implies that $Y(\cdot, x)$ is absolutely continuous $P_0-a.s.$, by (H.0)' and (H.2)'. In particular, $\{Y(t,x)\}_{0\le t\le 1}$ on $(\mathbf{R}, \mathbf{B}(\mathbf{R}), P_0)$ belongs to the set A.
For $n \ge 1$ and $\alpha \in (0,1)$, by Jensen's inequality and (H.1)',

$$
\begin{aligned}
\infty > \sup_{m\ge 1} \mathbf{e}_m \ge \mathbf{e}_n \quad &\ge \quad E_0[\int_0^{1-\alpha} ds \biggl(\frac{1}{\alpha}\int_s^{s+\alpha} h\biggl(\frac{dY_n(t,x)}{dt}\biggr)dt\biggr)] \\
&\ge \quad E_0[\int_0^{1-\alpha} h\biggl(\frac{Y_n(s+\alpha,x)-Y_n(s,x)}{\alpha}\biggr)ds].
\end{aligned} \tag{3.52}
$$

Let $n \to \infty$ and then $\alpha \to 0$ in (3.31). Then the proof of the first part is over by Fatou's lemma since $\sup_{m\ge 1} \mathbf{e}_m \le \mathbf{e}$.
The following together with Proposition 2 in section 4 completes the proof: by (2.12),

$$
Y(t,x) = Y(0,x) + \int_0^t \frac{\partial F_s^{-1}(F_s(Y(s,x)))}{\partial s}\, ds \quad (0 \le t \le 1) \quad P_0 - a.s..
$$

Q. E. D.

4 Appendix

In this section we solve SP positively for **R**-valued, time-inhomogeneous stochastic processes.

Proposition 2 *Suppose that (H.3)' holds, and that there exists $\{Y(t)\}_{0\le t\le 1}, \in A$, which satisfies*

$$
Y(t) = Y(0) + \int_0^t b^Y(s, Y(s))ds \quad (0 \le t \le 1) \ a.s. \tag{4.53}
$$

for some $b^Y(t,x) \in L^1([0,1]\times \mathbf{R} : \mathbf{R}, p(t,x)dtdx)$. Then the following holds:

$$
Y(t) = F_t^{-1}(F_0(Y(0))) \quad (t \in \mathbf{Q}\cap[0,1]) \quad a.s.. \tag{4.54}
$$

In particular, $\lim_{s\in\mathbf{Q}\cap[0,1], s\to t} F_s^{-1}(F_0(Y(0)))$ exists and is equal to $Y(t)$ for all $t \in [0,1]$ a.s..

Remark 7 *If F_0 is not continuous, then SP does not always have a positive answer. For example, put $Y(t) \equiv tY(\omega)$ for a **R**-valued random variable $Y(\omega)$ on a probability space. Then $dY(t)/dt = Y(t)/t$ for $t > 0$. But, of course, $Y(t)$ is not a function of t and $Y(0) \equiv 0$.*

(Proof of Proposition 2). It is easy to see that the following holds:

$$
F_t(Y(t)) = F_0(Y(0)) \quad (t \in [0,1]) \text{ a.s..} \tag{4.55}
$$

Indeed,

$$
\frac{\partial F_t(x)}{\partial t} = -b^Y(t,x)p(t,x), \quad dtdx - a.e.
$$

since $b^Y(t,x) = b(t,x)$, $p(t,x)dtdx - a.e.$, and henceforth by (H.3)'

$$\begin{aligned} & E[\sup_{0\le t\le 1} |F_t(Y(t)) - F_0(Y(0))|] \\ \le & \int_0^1 E[\left|\frac{\partial F_s(Y(s))}{\partial s} + p(s,Y(s))b^Y(s,Y(s))\right|]ds = 0. \end{aligned}$$

Since $\{Y(t)\}_{0\le t\le 1}$ is continuous, the proof is over by (4.3) and by the following:

$$P(F_t^{-1}(F_t(Y(t))) = Y(t)(t \in [0,1] \cap \mathbf{Q})) = 1. \tag{4.56}$$

We prove (4.4). For $(t,x) \in [0,1] \times \mathbf{R}$ for which $F_t(x) \in (0,1)$,

$$F_t^{-1}(F_t(x)) \le x,$$

and for $t \in [0,1]$, the set $\{x \in \mathbf{R} : F_t^{-1}(F_t(x)) < x, F_t(x) \in (0,1)\}$ can be written as a union of at most countably many disjoint intervals of the form $(a,b]$ for which $P(a < Y(t) \le b) = 0$, provided that it is not empty.
Indeed, if $F_t^{-1}(F_t(x)) < x$ and if $F_t(x) \in (0,1)$, then

$$\begin{aligned} & \{y \in \mathbf{R} : F_t^{-1}(F_t(y)) < y, F_t(y) = F_t(x)\} \\ = & (F_t^{-1}(F_t(x)), \sup\{y \in \mathbf{R} : F_t(y) = F_t(x)\}]. \end{aligned}$$

Q. E. D.

(Acknowledgement) We would like to thank Prof. M. Takeda for a useful discussion on Salisbury's problem.

References

[1] Breiman, L. (1992), *Probability*, SIAM, Philadelphia.

[2] Brenier, Y. and Benamou, J. D. (1999), *A numerical method for the optimal mass transport problem and related problems*, Contemporary Mathematics **226**, Amer. Math. Soc., Providence, 1-11.

[3] Brenier, Y. and Benamou, J. D. (2000), *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84**, 375-393.

[4] Carlen, E. A. (1984), *Conservative diffusions*, Commun. Math. Phys. **94**, 293-315.

[5] Carlen, E. A. (1986), *Existence and sample path properties of the diffusions in Nelson's stochastic machanics*, Lecture Notes in Mathematics **1158**, Springer, Berlin Heidelberg New York, 25-51.

[6] Çinlar, E. and Jacod, J. (1981), *Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures*, Seminar on Stochastic Processes 1981, Birkhauser, Boston Basel Berlin, 159-242.

[7] Dall'Aglio, G. (1991), *Frèchet classes: the beginning*, Mathematics and its applications **67**, Kluwer Academic Publishers, Dordrecht Boston London, 1-12.

[8] Doob, J. L. (1990), *Stochastic processes*, John Wiley & Sons, Inc., New York.

[9] Evans, L. C. (1998), *Partial differential equations*, AMS, Providence.

[10] Evans, L. C. and Gangbo, W. (1999), *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Mem. Amer. Math. Soc. **137**, No. 653.

[11] Fleming, W. H. and Soner, H. M. (1993), *Controlled Markov Processes and Viscosity Solutions*, Springer, Berlin Heidelberg New York.

[12] Gangbo, W and McCann, R. J. (1996), *The geometry of optimal transportation*, Acta Math. **177**, 113-161.

[13] Ikeda, N. and Watanabe, S. (1981), *Stochastic differential equations and diffusion processes*, North-Holland/Kodansha, Amsterdam New York Oxford Tokyo.

[14] Jordan, R., Kinderlehrer, D. and Otto, F. (1998), *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29**, 1-17.

[15] Mikami, T. (1990), *Variational processes from the weak forward equation*, Commun. Math. Phys. **135**, 19-40.

[16] Mikami, T. (2000), *Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric*, Appl. Math. Optim. **42**, 203-227.

[17] Nelsen, R. B. (1999), *An Introduction to Copulas*, Lecture Notes in Statistics **139**, Springer, Berlin Heidelberg New York.

[18] Otto, F. (2001), *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26**, 101-174.

[19] Quastel, J. and Varadhan, S. R. S. (1997), *Diffusion semigroups and diffusion processes corresponding to degenerate divergence form operators*, Comm. Pure Appl. Math. **50**, 667-706.

[20] Rachev, S. T. and Rüschendorf, L. (1998), *Mass transportation problems, Vol. I: Theory*, Springer, Berlin Heidelberg New York.

[21] Salisbury, T. S. (1986), *An increasing diffusion*, Seminar on Stochastic Processes 1984, Birkhauser, Boston Basel Berlin, 173-194.