Optimal control for absolutely continuous stochastic processes and the mass transportation problem

Toshio Mikami

Series #494. October 2000

HOKKAIDO UNIVERSITY
PREPRINT SERIES IN MATHEMATICS

#470 M. Nakamura and T. Ozawa, Small solutions to nonlinear wave equations in the Sobolev spaces, 27 pages. 1999.

#471 K. Ito and Y. Kohsaka, Stability of a stationary solution for evolving boundaries of symmetric three-phases driven by surface diffusion, 36 pages. 1999.

#472 Y. Giga and M.-H. Sato, A level set approach to semicontinuous viscosity solutions for Cauchy problems, 32 pages. 1999.

#473 R. Yoneda, Characterization of Bloch space and Besov spaces by oscillations, 31 pages. 1999.

#474 T. Nakazi, Norm inequalities for some singular integral operators, 13 pages. 1999.

#475 A. Inoue and Y. Kasahara, Asymptotics for prediction errors of stationary processes with reflection positivity, 15 pages. 1999.

#476 T. Sano, On affine parallels of generic plane curves, 8 pages. 1999.

#477 T. Nakazi, On an invariant subspace whose common zero set is the zeros of some function, 11 pages. 1999.

#478 M.-H. Giga and Y. Giga, Generalized motion by nonlocal curvature in the plane, 68 pages. 2000.

#479 M.-H. Giga and Y. Giga, Crystalline and level set flow – Convergence of a crystalline algorithm for a general anisotropic curvature flow in the plane, 16 pages. 2000.

#480 A. Arai and M. Hirokawa, Stability of ground states in sectors and its application to the Wigner-Weisskopf model, 16 pages. 2000.

#481 T. Nakazi, Two dimensional Q-algebras, 11 pages. 2000.

#482 N. H. Bingham and A. Inoue, Tauberian and Mercerian theorems for systems of kernels, 16 pages. 2000.

#483 N. H. Bingham and A. Inoue, Abelian, Tauberian and Mercerian theorems for arithmetic sums, 29 pages. 2000.

#484 I. A. Bogaevski and G. Ishikawa, Lagrange mappings of the first open Whitney umbrella, 22 pages. 2000.

#485 A. Arai and H. Kawano, A class of deformations of the Schrödinger representation of the Heisenberg commutation relation and exact solution to a Heisenberg equation and a Schrödinger equation, 22 pages. 2000.

#486 T. Nakazi, Functions in N_+ with the positive real parts on the boundary, 21 pages. 2000.

#487 Y. Shibukawa, Classification of the R-operator, 36 pages. 2000.

#488 A. Inoue, Asymptotic behaviour for partial autocorrelation functions of fractional ARIMA processes, 20 pages. 2000.

#489 S. Ohtani, Construction of unramified Galois extensions over maximal abelian extensions of algebraic number fields, 14 pages. 2000.

#490 T. Nakazi and T. Yamamoto, The real part of an outer function and a Helson-Szegö weight, 13 pages. 2000.

#491 A. Yamagami, On Gouvêás conjecture on controlling the conductor, 11 pages. 2000.

#492 I. Tsuda and M. Hatakeyama, Making sense of internal logic: Theory and a case study, 10 pages. 2000.

#493 I. Tsuda, Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems, 73 pages. 2000.

Optimal control for absolutely continuous stochastic processes and the mass transportation problem

Toshio Mikami
Department of Mathematics
Hokkaido University
Sapporo 060-0810, Japan
mikami@math.sci.hokudai.ac.jp

ABSTRACT

We study the optimal control problem for $\mathbf{R}^d$-valued absolutely continuous stochastic processes with given marginal distributions at every time. This can be considered as a generalization of Monge-Kantorovich problem in which they fix marginal distributions only at an initial and a terminal times. When $d = 1$, we show that there exists a minimizer which is a function of a time and an initial point, that is, an optimal mass transportation. When $d > 1$, we only show that minimizers satisfy the same ordinary differential equation.

0. Introduction.

Monge-Kantorovich problem (MKP) plays a crucial role in many fields and has been studied by many authors (see [5]-[7], [9], [12], [14]-[16], [20], [22], [23], [25], [26] 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 $\mathbf{R}^d$. As Monge's problem, consider the following: find the minimizer of

$$V_{h(\cdot)}(Q_0, Q_1) \equiv \inf\{\int_{\mathbf{R}^d} h(\varphi(x) - x)Q_0(dx) : Q_0\varphi^{-1} = Q_1\}. \qquad (0.1).$$

As Kantorovich's approach to (0.1), the following is known: find first the minimizer of

$$\inf\{\int_{\mathbf{R}^d\times\mathbf{R}^d} h(y-x)\mu(dxdy) : \mu(dx \times \mathbf{R}^d) = Q_0(dx), \mu(\mathbf{R}^d \times dy) = Q_1(dy)\}, \quad (0.2).$$

and then show that the minimizer has the form of $Q_0(dx)\delta_{\varphi(x)}(dy)$ for some Borel measurable function φ on $\mathbf{R}^d$.

Under the following condition, (0.1) has a unique minimizer.

(H.0). $Q_0(dx)$ is absolutely continuous with respect to dx. (0.2) is finite.

(H.1). $h : \mathbf{R}^d \mapsto [0, \infty)$ is strictly convex.
(H.2). For any $r > 0$ and $\theta \in (0, \pi)$, there exists $M > 0$ such that the following holds: for any $x \in \mathbf{R}^d$ for which $|x| \geq M$, there exists $y \in \mathbf{R}^d$ such that L takes a maximum at x over the set $\{z \in \mathbf{R}^d : |z - x||y| \cos(\theta/2) \leq < y, z - x > \leq r|y|\}$.
(H.3). $\lim_{|u| \to \infty} h(u)/|u| = \infty$.

Theorem 0.1. *(see [12, Theorem 1.2]). Suppose that (H.0)-(H.3) hold. Then (0.1) has a unique minimizer and is equal to (0.2).*

The infimum in (0.2) is equal to the following:

$$\inf\{E[\int_0^1 h(d\phi(t)/dt)dt] : \{\phi(t)\}_{0 \leq t \leq 1} \text{ is absolutely continuous, a.s.,} \quad (0.3).$$
$$P\phi(t)^{-1} = Q_t(t = 0, 1)\}$$

(see e.g. [11, p. 35, Example 8.1]), and the minimizer $\{\phi^o(t)\}_{0 \leq t \leq 1}$ is a function of t and $\phi^o(0)$ under (H.0)-(H.3). In this paper we use the same notation P for different probability measures, for the sake of simplicity, when it is not confusing.

Suppose that we are given a solution to the following PDE: for any infinitely differentiable function $f : \mathbf{R}^d \mapsto \mathbf{R}$ with a compact support and any $t \in [0, 1]$,

$$\int_{\mathbf{R}^d} f(x)p(t,x)dx - \int_{\mathbf{R}^d} f(x)p(0,x)dx = \int_0^t ds \int_{\mathbf{R}^d} < \nabla f(x), b(s,x) > p(s,x)dx,$$
$$\int_{\mathbf{R}^d} p(t,x)dx = 1,$$
$$p(t,x) \geq 0 \quad dx - \text{a.e.},$$
$$(0.4).$$

where $b(t,x) \in L^1([0,1] \times \mathbf{R}^d; \mathbf{R}^d : p(t,x)dxdt)$. Let $L(t,x;u) : [0,1] \times \mathbf{R}^d \times \mathbf{R}^d \mapsto [0, \infty)$ be convex in u. In this paper we study the minimizers of

$$\inf\{E[\int_0^1 L(t, \phi(t); d\phi(t)/dt)dt] : \{\phi(t)\}_{0 \leq t \leq 1} \text{ is absolutely continuous, a.s.,} \ (0.5).$$
$$P\phi(t)^{-1} = p(t,x)dx(0 \leq t \leq 1)\},$$

which can be considered as a generalization of MKP.

Remark 0.1. (0.4) is the conservation equation for the mass transportation on the probability space $(\mathbf{R}^d, \mathbf{B}(\mathbf{R}^d), P_0(dx) \equiv p(0,x)dx)$ by a nonrandom $\{X(t,x)\}_{0 \leq t \leq 1, x \in \mathbf{R}^d}$ for which

$$X(t,x) = x + \int_0^t b(s, X(s,x))ds, \qquad (0.6).$$

provided that it exists. But as examples of (0.4), we would like to point out the Fokker-Planck equation and the porous medium equation, etc. (see e.g. [8]), that is, $b(t,x)$ in (0.4) can depend on $p(t,x)$.

When (0.4) contains a linear second-order differential operator, Markov control problem that can be obtained from (0.5) with $\phi(t)$ and $d\phi(t)/dt$ replaced by continuous semimartingales and progressively measurable functions, respectively, was studied in [17] where the minimizer satisfies the Itô-type stochastic differential equation.

In this paper we show that the set over which the infimum is taken in (0.5) is not empty, and study the existence, the uniqueness and the time evolution of the minimizer of (0.5). We would also like to know if the minimizer is a function of a time and an initial point. This problem can be considered as a modification of Salisbury's problem (see [4], [10], [24]):

(SP) Is a continuous strong Markov process which is of bounded variation in time a function of an initial point and of 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 [4]). When $d > 1$, a counter example is known (see [24]).

Theorem 0.1 suggests us to consider SP in the framework of the variational problem:

(SPV) Does there exist a minimizer, of (0.5), which is a function of an initial point and of a time?

When $L(t,x;u) = |u|^2$ and $p(t,x)$ satisfies the Fokker-Planck equation with sufficiently smooth coefficients, the optimization problem (0.5) was considered in [18] 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 $L(t,x;u) = |u|^2$.

In this paper, in case where $d = 1$, we give positive answers to SP for time-inhomogeneous stochastic processes and to SPV (see Theorem 1.1, and Theorem 1.2 and Corollary 1.5, respectively). In particular, we show that an absolutely continuous stochastic process $\{\phi(t)\}_{0\le t\le 1}$ which satisfies an ordinary differential equation

(ODE) and for which $P\phi(t)^{-1}$ equals to $p(t,x)dx$ for all $t \in [0,1]$ is unique, by the representation formula of the stochastic process, from which one concludes that a minimizer, of (0.5), which satisfies an ODE is unique. (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..) Theorem 1.1 is a slight generalization of [4] 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.

Let $(\Omega, \mathbf{B}, P)$ be a probability space, and $\{\mathbf{B}_t\}_{t\ge 0}$ 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 $\{W(t)\}_{t\ge 0}$ denote a d-dimensional $(\mathbf{B}_t)$-Wiener process (see e.g. [11] or [13]). For $\varepsilon > 0$, a $\mathbf{R}^d$-valued $(\mathbf{B}_t)$-progressively measurable $\{u(t)\}_{0\le t\le 1}$, and $t \in [0,1]$, put

$$X^{\varepsilon,u}(t) = X_o + \int_0^t u(s)ds + \varepsilon W(t). \tag{0.7}$$

As $\varepsilon \to 0$, the infimum of

$$E[\int_0^1 L(t, X^{\varepsilon,u}(t); u(t))dt] \tag{0.8}$$

over all $\{u(t)\}_{0\le t\le 1}$ for which $PX^{\varepsilon,u}(t)^{-1}$ is equal to $p(t,x)dx$ for all $t \in [0,1]$ seems to converge to (0.5), which will be rigorously discussed in section 1 (see Theorem 1.3).

We also show that the minimizers of (0.5) exist and satisfy the same ODE when L is striclty convex in u (see Theorems 1.3 and 1.4). In particular, when $d = 1$, we show that the minimizer of (0.5) exists uniquely and is a function of a time and an initial point (see Corollary 1.5). In Corollary 1.6, we give the existence of an absolutely continuous stochastic process, with given marginal distributions, which satisfies an ODE (see [2], [3], [17], [21] and the references therein for the construction of a Markov diffusion process with given marginal distributions).

SPV for $d \ge 2$ is our future problem.

Our main tool in the proof is the weak convergence method, the theory of Copulas, and the result on the construction of a Markov diffusion process from a family of marginal distributions.

In section 1 we state our main result, for a cost function $L = h(u)$ when $d = 1$ and for $L(t,x;u)$ which grows at least of order of $|u|^2$ as $u \to \infty$ when $d \ge 1$, which will be proved in section 2.

We give notations in the following.

$A \equiv \{\{Y(t)\}_{0\le t\le 1} : \{Y(t)\}_{0\le t\le 1}$ is absolutely continuous, a.s. and $PY(t)^{-1} = p(t,x)dx$ for all $t \in [0,1]\}$.

$A_n \equiv \{\{Y(t)\}_{0\le t\le 1} : \{Y(t)\}_{0\le t\le 1}$ is absolutely continuous, a.s. and $PY(t)^{-1} = p(t,x)dx$ for all $t = 0, 1/n, \cdots, 1\}$ $(n \ge 1)$.

$R(h,\delta) \equiv \sup\{(L(t,x;u) - L(s,y;u))/(1 + L(s,y;u)) : |t-s| < h, |x-y| < \delta, u \in \mathbf{R}^d\}$ $(h,\ \delta > 0)$.

$\partial_u L(t,x;u) \equiv \{z \in \mathbf{R}^d : L(t,x;v) - L(t,x;u) \ge < z, v-u >$ for all $v \in \mathbf{R}^d\}$ $(t \in [0,1],\ x,\ u \in \mathbf{R}^d)$.

$A^\varepsilon \equiv \{\{u(t)\}_{0\le t\le 1} : PX^{\varepsilon,u}(t)^{-1} = p(t,x)dx (0 \le t \le 1)\}$.

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

$\mathbf{e}^0 \equiv \inf\{E[\int_0^1 L(t, Y(t); dY(t)/dt)dt] : \{Y(t)\}_{0\le t\le 1} \in A\}$.

$\tilde{A} \equiv \{\{B(t,x)\}_{(t,x)\in[0,1]\times\mathbf{R}^d} \subset \mathbf{R}^d :$ (0.4) with b replaced by B holds $\}$.

$\tilde{\mathbf{e}}^0 \equiv \inf\{\int_0^1 \int_{\mathbf{R}^d} L(t,y; B(t,y))p(t,y)dydt : B \in \tilde{A}\}$.

$\tilde{A}^\varepsilon \equiv \{\{B(t,x)\}_{(t,x)\in[0,1]\times\mathbf{R}^d} \subset \mathbf{R}^d :$ (0.4) with $< \nabla f(x), b(s,x) >$ replaced by $\varepsilon^2 \triangle f(x)/2 + < \nabla f(x), B(s,x) >$ holds$\}$.

$\tilde{\mathbf{e}}^\varepsilon \equiv \inf\{\int_0^1 \int_{\mathbf{R}^d} L(t,y; B(t,y))p(t,y)dydt : B \in \tilde{A}^\varepsilon\}$.

When $d = 1$, put

$\mathbf{e} \equiv \inf\{E[\int_0^1 h(dY(t)/dt)dt] : \{Y(t)\}_{0\le t\le 1} \in A\}$,

$\mathbf{e}_n \equiv \inf\{E[\int_0^1 h(dY(t)/dt)dt] : \{Y(t)\}_{0\le t\le 1} \in A_n\}$ $(n \ge 1)$,

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

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

1. Main result.

In this section we give our main result.

We first consider the one-dimensional case.

Under the following condition, SP can be solved positively even for $\mathbf{R}$-valued, time-inhomogeneous stochastic processes.

(H.0)'. $d = 1$ and $F_t(x)$ is differentiable with locally bounded partial derivatives on $[0,1] \times \mathbf{R}$.

Theorem 1.1. *Suppose that (H.0)' holds, and that there exists* $\{Y(t)\}_{0 \le t \le 1}, \in A$, *which satisfies*

$$Y(t) = Y(0) + \int_0^t b^Y(s, Y(s))ds \quad (0 \le t \le 1) \text{ a.s.} \tag{1.1}$$

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

$$Y(t) = \lim_{s \in \mathbf{Q} \cap [0,1], s \to t} F_s^{-1}(F_0(Y(0))) \quad (t \in [0,1]) \quad a.s.. \tag{1.2}$$

Remark 1.1. 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 $\mathbf{R}$-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 $Y(0) \equiv 0$.

Under the following hypothesis, by the theory of Copulas, we give a positive answer to SPV.

(H.0)'. $d = 1$, $\{\mathbf{e}_n\}_{n \ge 1}$ is bounded, and $F_t^{-1}(u)$ is absolutely continuous in $t \in [0,1]$, du-a.e..

(H.1)'. $h(u) = \tilde{h}(|u|)$, where $\tilde{h} : [0,\infty) \mapsto [0,\infty)$ is convex.

Theorem 1.2. *Suppose that (H.0)'-(H.1)' hold. Then the set A is not empty and* $\{F_t^{-1}(F_0(x))\}_{0 \le t \le 1}$ *on* $(\mathbf{R}^d, \mathbf{B}(\mathbf{R}^d), P_0)$ *is a minimizer of* $\mathbf{e}$. *Moreover if (H.0)' holds, then* $\{F_t^{-1}(F_0(x))\}_{0 \le t \le 1}$ *is the unique minimizer, of* $\mathbf{e}$, *that satisfies an ODE.*

Remark 1.2. $\{X(t) \equiv F_t^{-1}(F_0(x))\}_{0 \le t \le 1}$ in Theorem 1.2 satisfies the following: for any t and s for which $0 \le s \le t \le 1$,

$$V_{h(\cdot)}(p(s,x)dx, p(t,x)dx) = E_0[h(X(t) - X(s))], \tag{1.3}$$

since

$$X(t) = F_t^{-1}(F_s(X(s))) \tag{1.4}$$

(see e.g. [23, Chap. 3.1]).

For $d \geq 1$, we study the minimizers of $\mathbf{e}$ for a cost function $L(t,x;u)$, under the assumption on the square integrability of b for which (0.4) holds.

Let us state assumptions before we state the result.

(H.0)". $\tilde{\mathbf{e}}^0$ is finite.

(H.1)". $L(t,x;u) : [0,1] \times \mathbf{R}^d \times \mathbf{R}^d \mapsto [0,\infty)$ is convex in u, and $R(h,\delta) \to 0$ as $(h,\delta) \to (0,0)$.

(H.3)'. There exists $q \geq 2$ such that the following holds:

$$0 < \liminf_{|u|\to\infty}(\inf\{L(t,x;u) : (t,x) \in [0,1] \times \mathbf{R}^d\})/|u|^q, \tag{1.5}$$

$$\sup\{\sup_{z\in\partial_u L(t,x;u)} |z|/(1+|u|)^{q-1} : (t,x,u) \in [0,1] \times \mathbf{R}^d \times \mathbf{R}^d\} \equiv C_{\nabla L} < \infty. \tag{1.6}$$

(H.4). $p(t,\cdot)$ is absolutely continuous dt-a.e., and for q in (H.3)',

$$\int_0^1 \int_{\mathbf{R}^d} |\nabla_x p(t,x)/p(t,x)|^q p(t,x)dxdt < \infty. \tag{1.7}$$

Remark 1.3. (H.1)" implies the continuity of $L(\cdot,\cdot;u)$ for each $u \in \mathbf{R}^d$. We need (H.4) to make use of the result on the construction of a Markov diffusion process of which the marginal distributions are $\{p(t,x)dx\}_{0\leq t\leq 1}$ in (0.4). (1.7) holds if $b(t,x)$ in (0.4) 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:

$$\int_{\mathbf{R}^d} |\nabla_x p(0,x)/p(0,x)|^q p(0,x)dx < \infty. \tag{1.8}$$

The following theorem implies the existence of a minimizer of $\mathbf{e}^0$.

Theorem 1.3. *Suppose that (H.0)", (H.1)", (H.3)' and (H.4) hold. Then the sets A^ε $(\varepsilon > 0)$ and A are not empty, and the following holds:*

$$\lim_{\varepsilon\to 0} \mathbf{e}^\varepsilon = \mathbf{e}^0 = \tilde{\mathbf{e}}^0. \tag{1.9}$$

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

$$\lim_{\varepsilon\to 0} E[\int_0^1 L(t, X^{\varepsilon,u^\varepsilon}(t); u^\varepsilon(t))dt] = \mathbf{e}^0, \qquad (1.10).$$

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

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

Theorem 1.4. *Suppose that (H.0)", (H.1)", (H.3)' and (H.4) hold. Then for any minimizer $\{X(t)\}_{0\le t\le 1}$ of $\mathbf{e}^0$, $b^X(t,x) \equiv E[dX(t)/dt|(t, X(t) = x)]$ is a minimizer of $\tilde{\mathbf{e}}^0$. In particular, if 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 $\mathbf{e}^0$,*

$$dX(t)/dt = b^o(t, X(t)) \quad dtdP - a.e.. \qquad (1.11).$$

By Theorems 1.1, 1.3 and 1.4, we obtain the following.

Corollary 1.5. *Suppose that (H.0)', (H.0)", (H.1)", (H.3)' and (H.4) hold, and that L is strictly convex in u. Then $\mathbf{e}^0$ has the unique minimizer $\{\lim_{s\in\mathbf{Q}\cap[0,1],s\to t} F_s^{-1}(F_0(X(0)))\}_{0\le t\le 1}$.*

By Theorems 1.3 and 1.4, we easily obtain the following.

Corollary 1.6. *Suppose that (H.0)" with $L = |u|^2$ and (H.4) with $q = 2$ hold. Then there exists a stochastic process $\{X(t)\}_{0\le t\le 1}, \in A$, which satisfies an ODE.*

The following is an example of the minimizer of $\tilde{\mathbf{e}}^0$.

Proposition 1.7. *Suppose that $L = |u|^2$, and that (H.0)" and (H.4) hold, and that for any $M > 0$,*

$$ess.inf\{p(t,x) : t \in [0,1], |x| \le M\} > 0. \qquad (1.12).$$

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

2. Proof of the result.

In this section we prove the result given in section 1.

(Proof of Theorem 1.1). The following is true:

$$F_t(Y(t)) = F_0(Y(0)) \quad (t \in [0,1]) \text{ a.s.}, \tag{2.1}$$

since, by the assumption, (0.4) with b replaced by b^Y holds, from which we obtain

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

and henceforth

$$E[\sup_{0\le t\le 1} |F_t(Y(t)) - F_0(Y(0))|]$$
$$\le \int_0^1 E[|\partial F_s(Y(s))/\partial s + p(s,Y(s))b^Y(s,Y(s))|]ds = 0.$$

Here we used the absolute continuity, of $\{F_t(Y(t))\}_{0\le t\le 1}$ in t, which can be obtained by the assumption.

Since $\{Y(t)\}_{0\le t\le 1}$ is continuous, the proof is over by (2.1) and by the following:

$$P(F_t^{-1}(F_t(Y(t))) = Y(t)(t \in [0,1] \cap \mathbf{Q})) = 1. \tag{2.2}$$

(2.2) is true. Indeed, 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. This is true, since if $F_t^{-1}(F_t(x)) < x$ and if $F_t(x) \in (0,1)$, then there exist a and b for which $a \le x \le b$ $(a < b)$ and such that $F_t(a) = F_t(b)$. This implies $P(a < Y(t) \le b) = 0$, since $PY(t)^{-1} = p(t,x)dx$.

Q. E. D.

(Proof of Theorem 1.2). Put for $t \in [0,1]$, $x \in \mathbf{R}$ and $n \ge 1$,

$$Y(t,x) = F_t^{-1}(F_0(x)), \tag{2.3}$$
$$Y^n(t,x) = Y([nt]/n, x) + n(t - [nt]/n)(Y(([nt]+1)/n, x) - Y([nt]/n, x)). \tag{2.4}$$

Then $\{Y(t,x)\}_{0\le t\le 1}$ on $(\mathbf{R}, \mathbf{B}(\mathbf{R}), P_0)$ belongs to the set A, and the following holds (see [23, Chap. 3.1], [6], or [19]): for $n \ge 1$ and $\alpha \in (0,1)$, by the discussion in (0.3) and (1.3)-(1.4),

$$\begin{aligned}\infty > \sup_{m\ge 1} \mathbf{e}_m \ge \mathbf{e}_n &= \sum_{k=1}^{n} V_{h(n\cdot)}(p((k-1)/n, x)dx, p(k/n,x)dx)/n \qquad (2.5).\\ &= E_0[\int_0^1 h(dY^n(t,x)/dt)dt]\\ &\ge E_0[\int_0^{1-\alpha} ds \int_s^{s+\alpha} h(dY^n(t,x)/dt)dt/\alpha]\\ &\ge E_0[\int_0^{1-\alpha} dsh((Y^n(s+\alpha,x) - Y^n(s,x))/\alpha)]\end{aligned}$$

by Jensen's inequality (see e.g. [1]). (Notice that h is convex by (H.1)'.) Let $n \to \infty$ and then $\alpha \to 0$ in (2.5). Then the proof of the first part is over by Fatou's lemma, since

$$\lim_{n\to\infty} Y^n(t,x) = Y(t,x) \quad (0 \le t \le 1), P_0(dx) - a.s.,$$

and since $\sup_{m\ge 1} \mathbf{e}_m \le \mathbf{e}$.

The following together with Theorem 1.1 completes the proof:

$$dY(t,x)/dt = \partial F_t^{-1}(F_t(Y(t,x)))/\partial t,$$

since $F_t(F_t^{-1}(u)) = u$ $(0 < u < 1)$ (see e.g. [19]).

Q. E. D.

Before we give the proof of Theorem 1.3, let us state and prove three technical lemmas.

Lemma 2.1. *Suppose that (H.0)", (H.3)' and (H.4) hold. Then for any* $B \in \tilde{A}$ *for which* $\int_0^1 \int_{\mathbf{R}^d} L(t,x;B(t,x))p(t,x)dxdt$ *is finite and for any* $\varepsilon > 0$, *there exists a Markov process* $\{Z^{\varepsilon,B}(t)\}_{0\le t\le 1}$ *such that the following holds:*

$$\begin{aligned} Z^{\varepsilon,B}(t) =& X_o + \int_0^t (\varepsilon^2 \nabla p(s, Z^{\varepsilon,B}(s))/(2p(s, Z^{\varepsilon,B}(s))) \qquad (2.6).\\ &+ B(s, Z^{\varepsilon,B}(s)))ds + \varepsilon W(t),\\ PZ^{\varepsilon,B}(t)^{-1} =& p(t,x)dx \quad (0 \le t \le 1). \qquad (2.7).\end{aligned}$$

In particular, $\{\varepsilon^2\nabla p(t,Z^{\varepsilon,B}(t))/(2p(t,Z^{\varepsilon,B}(t)))+B(t,Z^{\varepsilon,B}(t))\}_{0\le t\le 1}$ belongs to the set A^ε.

(Proof). The proof is done by the following (see [2, 3] and also [17, 21]): for any infinitely differentiable function $f:\mathbf{R}^d\mapsto\mathbf{R}$ with a compact support and any $t\in[0,1]$,

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

and by (H.3)' and (H.4),

$$\int_0^1\int_{\mathbf{R}^d}|\nabla_x p(t,x)/(2p(t,x))+B(t,x)|^2 p(t,x)dxdt<\infty.$$

Q. E. D.

The following lemma can be shown by the standard argument and the proof is omitted.

Lemma 2.2. *Any $\{u^\varepsilon(t)\}_{0\le t\le 1},\in A^\varepsilon$ $(\varepsilon>0)$, for which $E[\int_0^1|u^\varepsilon(t)|^2dt]$ is bounded is tight in $C([0,1];\mathbf{R}^d)$.*

Lemma 2.3. *For any weakly convergent sequence $\{X^n(t)\equiv X^{\varepsilon_n,u^{\varepsilon_n}}(t)\}_{0\le t\le 1}\in A^{\varepsilon_n}$ $(n\ge 1)$ $(\varepsilon_n\to 0$ as $n\to\infty)$ for which $E[\int_0^1|u^{\varepsilon_n}(t)|^2dt]$ is bounded, the weak limit point $\{X(t)\}_{0\le t\le 1}$ in $C([0,1];\mathbf{R}^d)$ is absolutely continuous.*

(Proof). $\{X(t)\}_{0\le t\le 1}$ is of bounded variation in time. In fact, the following holds:

$$E[\sup_{m\ge 2}(\sup\{\sum_{i=1}^{m-1}|X(s_{i+1})-X(s_i)|;0\le s_1<\cdots<s_m\le 1\})]<\infty. \qquad (2.8).$$

This can be proved by the monotone convergence theorem, since for any $m\ge 2$, by the continuity of $\{X(t)\}_{0\le t\le 1}$,

$$\begin{aligned}&\sup\{\sum_{i=1}^{m-1}|X(s_{i+1})-X(s_i)|;0\le s_1<\cdots<s_m\le 1\}\\ &=\sup\{\sum_{i=1}^{m-1}|X(s_{i+1})-X(s_i)|;0\le s_1<\cdots<s_m\le 1, s_i\in\mathbf{Q}(i=1,\cdots,m)\},\end{aligned}$$

and since for any $m \geq 2$ and any $0 \leq s_i < t_i \leq s_{i+1} < t_{i+1} \leq 1$ $(i = 1, \cdots, m-1)$, by Jensen's inequality,

$$
\begin{aligned}
&E[(\sum_{i=1}^{m} |X(t_i) - X(s_i)|)^2] && (2.9).\\
&\leq (\sum_{i=1}^{m} |t_i - s_i|) E[\sum_{i=1}^{m} |(X(t_i) - X(s_i))/(t_i - s_i)|^2](t_i - s_i)\\
&\leq (\sum_{i=1}^{m} |t_i - s_i|) \liminf_{n\to\infty} E[\int_0^1 |u^{\varepsilon_n}(t)|^2 dt].
\end{aligned}
$$

In particular, by (2.9), the following holds with probability one: for any sequence $\{(s_i^m, t_i^m)\}_{m,i\geq 1} \subset \mathbf{Q} \times \mathbf{Q}$ for which $0 \leq s_i^m < t_i^m \leq s_{i+1}^m < t_{i+1}^m \leq 1$ and for which $\lim_{m\to\infty} \sum_{i\geq 1}(t_i^m - s_i^m) = 0$,

$$
\liminf_{m\to\infty} \sum_{i=1}^{\infty} |X(t_i^m) - X(s_i^m)| = 0. \qquad (2.10).
$$

This implies the absolute continuity of $\{X(t)\}_{0\leq t\leq 1}$.

Q. E. D.

Let us prove Theorem 1.3 by Lemmas 2.1-2.3.

(Proof of Theorem 1.3). By Lemma 2.1, the set A^ε is not empty for $\varepsilon > 0$. The proof of (1.9) is divided into the following two inequalities:

$$
\begin{aligned}
\limsup_{\varepsilon\to 0} \mathbf{e}^\varepsilon &\leq \tilde{\mathbf{e}}^0, && (2.11).\\
\liminf_{\varepsilon\to 0} \mathbf{e}^\varepsilon &\geq \mathbf{e}^0. && (2.12).
\end{aligned}
$$

We first prove (2.11). For $B \in \tilde{A}$ for which $\int_0^1 \int_{\mathbf{R}^d} L(t,x;B(t,x))p(t,x)dxdt$ is finite and for $\varepsilon > 0$, take a Markov process $\{Z^{\varepsilon,B}(t)\}_{0\leq t\leq 1}$ in Lemma 2.1. Then by (H.3)',

$$
\begin{aligned}
\mathbf{e}^\varepsilon &\leq E[\int_0^1 L(t, Z^{\varepsilon,B}(t); \varepsilon^2\nabla p(t, Z^{\varepsilon,B}(t))/(2p(t, Z^{\varepsilon,B}(t))) + B(t, Z^{\varepsilon,B}(t)))dt] \quad (2.13).\\
&\leq E[\int_0^1 L(t, Z^{\varepsilon,B}(t); B(t, Z^{\varepsilon,B}(t)))dt]
\end{aligned}
$$

$$+E[\int_0^1 C_{\nabla L}(|\varepsilon^2\nabla p(t,Z^{\varepsilon,B}(t))/(2p(t,Z^{\varepsilon,B}(t)))+B(t,Z^{\varepsilon,B}(t))|+1)^{q-1}$$
$$\times|\varepsilon^2\nabla p(t,Z^{\varepsilon,B}(t))/(2p(t,Z^{\varepsilon,B}(t)))|dt]$$
$$\to\int_0^1\int_{\mathbf{R}^d} L(t,x;B(t,x))p(t,x)dxdt \quad (\text{as } \varepsilon\to 0),$$

where we used the following in the last part of (2.13):

$$(q-1)/q+1/q=1.$$

To prove (2.12), by Lemmas 2.2-2.3, we only have to show the following: for any weakly convergent sequence $\{X^n(t)\equiv X^{\varepsilon_n,u^{\varepsilon_n}}(t)\}_{0\le t\le 1}\in A^{\varepsilon_n}$ $(n\ge 1)$ $(\varepsilon_n\to 0$ as $n\to\infty)$ for which $E[\int_0^1 L(t,X^n(t);u^{\varepsilon_n}(t))dt]$ is bounded and the weak limit point $\{X(t)\}_{0\le t\le 1}$ in $C([0,1];\mathbf{R}^d)$,

$$\liminf_{n\to\infty} E[\int_0^1 L(t,X^n(t);u^{\varepsilon_n}(t))dt]\ge E[\int_0^1 L(t,X(t);dX(t)/dt)dt]\ge \mathbf{e}^0. \quad (2.14)$$

(Notice that $\{X(t)\}_{0\le t\le 1}\in A$ by Lemma 2.3.)

(2.14) is true, since, in the same way as in inequalities of (2.5), by (H.1)" and Jensen's inequality, for $\alpha\in(0,1)$,

$$E[\int_0^1 L(t,X^n(t);u^{\varepsilon_n}(t))dt] \quad (2.15).$$
$$\ge E[\int_0^{1-\alpha} dsL(s,X^n(s);\int_s^{s+\alpha}u^{\varepsilon_n}(t)dt/\alpha)/(1+R(\alpha,\delta))$$
$$;\sup_{0\le t,s\le 1,|t-s|<\alpha}|X^n(t)-X^n(s)|<\delta]-R(\alpha,\delta)$$
$$\ge E[\int_0^{1-\alpha} L(s,X^n(s);(X^n(s+\alpha)-X^n(s))/\alpha)ds$$
$$;\sup_{0\le t,s\le 1,|t-s|<\alpha}|X^n(t)-X^n(s)|<\delta]/(1+R(\alpha,\delta))$$
$$-E[\int_0^{1-\alpha} C_{\nabla L}(|(X^n(s+\alpha)-X^n(s))/\alpha|+1)^{q-1}$$
$$\times|\varepsilon_n(W(s+\alpha)-W(s))/\alpha|ds]-R(\alpha,\delta).$$

Let $n\to\infty$ and then $\alpha\to 0$ and $\delta\to 0$ in (2.15). Then the proof is over by (H.1)", (H.3)', Skorohod's theorem (see e.g. [13]) and Fatou's lemma. Indeed, by taking a different probability space, one can assume that $\{\{X^n(t)\}_{0\le t\le 1}\}_{n\ge 1}$ converges, as $n \to \infty$, to $\{X(t)\}_{0\le t\le 1}$ in sup norm, a.s., and the following is true: for any $\beta \in (0, \delta/3)$,

$$E[\int_0^{1-\alpha} L(s, X^n(s); (X^n(s+\alpha) - X^n(s))/\alpha)ds; \sup_{0\le t,s\le 1, |t-s|<\alpha} |X^n(t) - X^n(s)| < \delta]$$
$$\ge E[\int_0^{1-\alpha} L(s, X(s); (X^n(s+\alpha) - X^n(s))/\alpha)ds; \sup_{0\le t,s\le 1, |t-s|<\alpha} |X(t) - X(s)| < \beta$$
$$, \sup_{0\le t\le 1} |X(t) - X^n(t)| < \beta]/(1 + R(0,\beta)) - R(0,\beta).$$

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.

(H.0)", (1.9), Lemmas 2.2-2.3, and (2.14) complete the proof of the rest.

Q. E. D.

(Proof of Theorem 1.4). $b^X(t,x)$ is a minimizer of $\tilde{\mathbf{e}}^0$ by (1.9), since for any infinitely differentiable function $f : \mathbf{R}^d \mapsto \mathbf{R}$ with a compact support and any $t \in [0,1]$,

$$\begin{aligned} E[f(X(t))] - E[f(X(0))] &= \int_0^t E[< \nabla f(X(s)), dX(s)/ds >]ds \\ &= \int_0^t E[< \nabla f(X(s)), E[dX(s)/ds|(s, X(s))] >]ds, \end{aligned}$$

which implies that $b^X(t,x) \in \tilde{A}$, and since by Jensen's inequality,

$$\mathbf{e}^0 = E[\int_0^1 L(t, X(t); dX(t)/dt)dt] \ge E[\int_0^1 L(t, X(t); b^X(t, X(t)))dt] \ge \tilde{\mathbf{e}}^0. \quad (2.16).$$

Next we prove the uniqueness of the minimizer of $\tilde{\mathbf{e}}^0$. Suppose that $\tilde{b}(t,x)$ is also a minimizer of $\tilde{\mathbf{e}}^0$. Then for any $\lambda \in (0,1)$, $\lambda b^X + (1-\lambda)\tilde{b} \in \tilde{A}$, and

$$\begin{aligned} &\tilde{\mathbf{e}}^0 \qquad\qquad (2.17). \\ &\le \int_0^1 \int_{\mathbf{R}^d} L(t, y; \lambda b^X(t,y) + (1-\lambda)\tilde{b}(t,y))p(t,y)dydt \\ &\le \lambda \int_0^1 \int_{\mathbf{R}^d} L(t, y; b^X(t,y))p(t,y)dydt + (1-\lambda) \int_0^1 \int_{\mathbf{R}^d} L(t, y; \tilde{b}(t,y))p(t,y)dydt = \tilde{\mathbf{e}}^0. \end{aligned}$$

By the strict convexity of L in u, $b^X(t,x) = \tilde{b}(t,x)$, $p(t,x)dxdt - a.e.$.

Finally we prove (1.11). Since L is strictly convex in u, the following holds by (1.9) and (2.16):

$$dX(t)/dt = b^X(t, X(t)) \quad dtdP - a.e.. \tag{2.18}$$

Hence

$$\begin{aligned} &E[\sup_{0\le t\le 1} |X(t) - X(0) - \int_0^t b^o(s, X(s))ds|] \\ &\le \int_0^1 E[|b^X(s, X(s)) - b^o(s, X(s))|]ds = 0 \end{aligned} \tag{2.19}$$

because of the uniqueness of the minimizer of $\tilde{\mathbf{e}}^0$.

Q. E. D.

(Proof of Proposition 1.7). From [17], $\tilde{\mathbf{e}}^\varepsilon = \mathbf{e}^\varepsilon$, and the minimizer of $\tilde{\mathbf{e}}^\varepsilon$ can be written as $\nabla_x \Phi^\varepsilon(t,x)$, where $\Phi^\varepsilon(t,\cdot) \in H^1_{loc}(\mathbf{R}^d)$ dt-a.e.. Since $\{\nabla_x\Phi^\varepsilon\}_{0<\varepsilon<1}$ is strongly bounded in $L^2([0,1]\times\mathbf{R}^d; p(t,x)dxdt)$ by (1.9), it is weakly compact in $L^2([0,1]\times\mathbf{R}^d; p(t,x)dxdt)$ (see [8, p. 639]). Let us denote a weak limit point by Ψ. Then Ψ is the unique minimizer of $\tilde{\mathbf{e}}^0$ by (1.9) and Theorem 1.4, since $\Psi \in \tilde{A}$, and since

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

In particular, $\{\nabla_x\Phi^\varepsilon\}_{0<\varepsilon<1}$ converges, as $\varepsilon \to 0$, to Ψ, strongly in $L^2([0,1]\times\mathbf{R}^d; p(t,x)dxdt)$, which completes the proof in the same way as in [17, Proposition 3.1].

Q. E. D.

Remark 2.1. If $V(t,x)$ and $p(t,x)$ in the proof of Porposition 1.7 is sufficiently smooth, then

$$\nabla_x\Phi^\varepsilon(t,x) = \nabla_x V(t,x) + \varepsilon^2\nabla_x p(t,x)/(2p(t,x)) \tag{2.20}$$

(see [18, section 1]).

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

References.

[1] Breiman L (1992) Probability. Classics in applied mathematics, Vol. 7. SIAM, Philadelphia

[2] Carlen EA (1984) Conservative diffusions. Commun. Math. Phys. 94:293-315.

[3] Carlen EA (1986) Existence and sample path properties of the diffusions in Nelson's stochastic machanics. In Stochastic processes-Mathematics and Physics Proceedings, Bielefeld 1984 (Albeverio S, Blanchard Ph, Streit L, eds.), Lecture Notes in Math., Vol. 1158, pp. 25-51. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

[4] Çinlar E, Jacod J (1981) Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures. Seminar on Stochastic Processes 1981 (Cinlar E, Chung KL, Getoor RK, eds.), pp. 159–242. Birkhäuser , Boston.

[5] Cuesta JA, Tuero-Diaz A (1993) A characterization for the solution of the Monge-Kantorovich mass transference problem. Statist. Probab. Lett. 16:147-152.

[6] Dall'Aglio G (1991) Frèchet classes: the beginning. In Advances in Probability Distributions with Given Marginals (Dall'Aglio G, Kotz S, Salinetti G, eds.), Mathematics and its applications, Vol. 67, pp. 1-12. Kluwer Academic Publishers, Dordrecht, Boston, London.

[7] Evans LC (1997) Partial differential equations and Monge-Kantorovich mass transfer. preprint.

[8] Evans LC (1998) Partial differential equations. Graduate Studies in Mathematics Vol. 19, AMS, Providence, RI, USA

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

[10] Fitzsimmons PJ (1999) Brownian space-time functions of zero quadratic variation depend only on time. Proceedings of AMS 127:2423-2429.

[11] Fleming WH, Soner HM (1993) Controlled Markov Processes and Viscosity Solutions. Springer - Verlag, Berlin, Heidelberg, New York, Tokyo

[12] Gangbo W, McCann RJ (1996) The geometry of optimal transportation. Acta Math. 177:113-161.

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

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

[15] Katz BS(editor) (1989) Nobel Laureates in Economic Sciences: A Biographical Dictionary. Garland Publishing Inc., New York.

[16] Levin VL (1990) General Monge-Kantorovich problem and its applications in measure theory and mathematical economics. In Functional Analysis, Optimization, and Mathematical Economics (Leifman LJ, ed.), pp. 141-176. Oxford Univ. Press, New York.

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

[18] Mikami T (1999) Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric. To appear in Appl. Math. Optim..

[19] Nelsen RB (1999) An Introduction to Copulas. Lecture Notes in Statis., Vol. 139, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

[20] Otto F, Villani C (2000) Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. 173: 361-400.

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

[22] Rachev ST (1984) The Monge-Kantorovich mass transference problem and its stochastic applications. Theory Probab. Appl. 29:647-676.

[23] Rachev ST, Rüschendorf L (1999) Mass transportation problems, Vol. I: Theory. Springer - Verlag, Berlin, Heidelberg, New York, Tokyo.

[24] Salisbury TS (1986) An increasing diffusion. Seminar on Stochastic Processes 1984 (Cinlar E, Chung KL, Getoor RK, eds.), pp. 173-194. Birkhäuser, Boston.

[25] Sudakov VN (1979) Geometric problems in the theory of infinite dimensional probability distributions. Proc. Steklov Inst. Math. 141:1-178.

[26] Talagrand M (1996) Transportation cost for Gaussian and other product measures. Geom. Funct. Anal. 6:587-600.