Monge-Kantorovitch Measure Transportation, Monge-Ampère Equation and the Itô Calculus

Denis Feyel and Ali Süleyman Üstünel

Dedicated to Professor Kiyosi Itô for his 88th birthday

Abstract.

Let (W, μ, H) be an abstract Wiener space assume two ν_i, $i = 1, 2$ probabilities on $(W, \mathcal{B}(W))$. Assume that the Wasserstein distance between ν_1 and ν_2 with respect to the Cameron-Martin norm

$$d_H(\nu_1, \nu_2) = \left\{ \inf_{\beta} \int_{W \times W} |x - y|_H^2 d\beta(x, y) \right\}^{1/2}$$

is finite, where the infimum is taken on the set of probability measures β on $W \times W$ whose first and second marginals are ν_1 and ν_2 and that ν_1 has regular disintegration along a sequence of finite dimensional projections. Then there exists a unique (cyclically monotone) map $T = I_W + \xi$, with $\xi : W \to H$, such that T maps ν_1 to ν_2 and the measure $\gamma = (I \times T)\nu_1$ is the unique solution of the Monge-Kantorovitch problem. Besides, if $\nu_2 \ll \mu^{1)}$, then T is stochastically invertible, i.e., there exists $S : W \to W$ such that $S \circ T = I_W$ ν_1 a.s. and $T \circ S = I_W$ ν_2 a.s. If $\nu_1 = \mu$, then there exists a 1-convex function ϕ in the Gaussian Sobolev space $\mathbb{D}_{2,1}$, such that $\xi = \nabla \phi$. These results imply that the quasi-invariant transformations of the Wiener space with finite Wasserstein distance from μ can be written as the composition of a transport map T and a rotation, i.e., a measure preserving map. We give also 1-convex sub-solutions using by calculating the Gaussian jacobian. Finally the full solutions of the Monge-Ampère equation on W are given with the help of the Itô calculus.

§1. Introduction

In 1781, Gaspard Monge has published his celebrated memoire about the most economical way of earth-moving [20]. The configurations of

[1] In fact this hypothesis is too strong, cf. Theorem 4.1.

excavated earth and remblai were modelized as two measures of equal
mass, say ρ and ν, that Monge had supposed absolutely continuous with
respect to the volume measure. Later Ampère has studied an analogous
question about the electricity current in a media with varying conductiv-
ity. In modern language of measure theory we can express the problem
in the following terms: let W be a Polish space on which are given two
positive measures ρ and ν, of finite, equal mass. Let $c(x, y)$ be a cost
function on $W \times W$, which is, usually, assumed to be positive. Does
there exist a map $T : W \to W$ such that $T\rho = \nu$ and T minimizes the
integral

$$\int_W c(x, T(x)) d\rho(x)$$

between all such maps? The problem has been further studied by Appell
[1, 2] and by Kantorovitch [16]. Kantorovitch has succeeded to trans-
form this highly nonlinear problem of Monge into a linear problem by
replacing the search for T with the search of a measure γ on $W \times W$
with marginals ρ and ν such that the integral

$$\int_{W \times W} c(x, y) d\gamma(x, y)$$

is the minimum of all the integrals

$$\int_{W \times W} c(x, y) d\beta(x, y)$$

where β runs in the set of probability measures on $W \times W$ whose
marginals are ρ and ν. Since then the problem addressed above is called
the Monge problem and the quest of the optimal measure is called the
Monge-Kantorovitch problem.

In this paper we survey and complete the recent results (cf. [12, 13])
about the Monge-Kantorovitch and the Monge problem in the frame
of an abstract Wiener space with a (infinitesimal) cost function c on
$W \times W$, which is singular with respect to the natural Fréchet topology
of this space. Let us explain all this more rigourously: let W be a
separable Fréchet space with its Borel sigma algebra $\mathcal{B}(W)$ and assume
that there is a separable Hilbert space H which is injected densely and
continuously into W, thus the topology of H is, in general, stronger than
the topology induced by W. The cost function $c : W \times W \to \mathbb{R}_+ \cup \{\infty\}$
is defined as

$$c(x, y) = |x - y|_H^2,$$

we suppose that $c(x, y) = \infty$ if $x - y$ does not belong to H. Clearly,
this choice of the function c is not arbitrary, in fact it is closely related

to Itô Calculus, hence also to the problems originating from Physics, quantum chemistry, large deviations, etc. Since for all the interesting measures on W, the Cameron-Martin space is a negligeable set, the cost function will be infinity very frequently. Let $\Sigma(\rho, \nu)$ denote the set of probability measures on $W \times W$ with given marginals ρ and ν. It is a convex, compact set under the weak topology $\sigma(\Sigma, C_b(W \times W))$. As explained above, the problem of Monge consists of finding a measurable map $T : W \to W$, called the optimal transport of ρ to ν, i.e., $T\rho = \nu$[2]) which minimizes the total cost

$$U \to \int_W |x - U(x)|_H^2 \, d\rho(x) \,,$$

between all the maps $U : W \to W$ such that $U\rho = \nu$. The Monge-Kantorovitch problem will consist of finding a measure on $W \times W$, which minimizes the function $\theta \to J(\theta)$, defined by

$$(1.1) \qquad J(\theta) = \int_{W \times W} |x - y|_H^2 \, d\theta(x, y) \,,$$

where θ runs in $\Sigma(\rho, \nu)$. Note that $\inf\{J(\theta) : \theta \in \Sigma(\rho, \nu)\}$ is the square of Wasserstein metric $d_H(\rho, \nu)$ with respect to the Cameron-Martin space H.

Any solution γ of the Monge-Kantorovitch problem will give a solution to the Monge problem provided that its support is included in the graph of a map. Hence our work consists of realizing this program. Although in the finite dimensional case this problem is well-studied in the path-breaking papers of Brenier [4] and McCann [18, 19], cf. also [25, 26], the things do not come up easily in our setting and the difficulty is due to the fact that the cost function is not continuous with respect to the Fréchet topology of W, for instance the weak convergence of the probability measures does not imply the convergence of the integrals of the cost function. In other words the function $|x - y|_H^2$ takes the value plus infinity "very often". On the other hand the results we obtain seem to have important applications to several problems of stochastic analysis that we shall explain while enumerating the contents of the paper.

In Section 2, are given the basic results of functional analysis on the Wiener space (cf., for instance [10, 28]) and the related probabilistic theory of convex functions developed in [11]. Section 3 deals with the derivation of some inequalities which control the Wasserstein distance.

---

[2]) We denote the push-forward of ρ by T, i.e., the image of ρ under T, by $T\rho$.

In particular, with the help of the Girsanov theorem and the Itô calculus, we give a very simple proof of an inequality, initially discovered by Talagrand ([24]), some simple applications are also illustrated. The facility with which one obtains these results gives an idea about the efficiency of the infinite dimensional techniques, namely the Itô calculus for the Monge-Kantorovitch like problems.

In Section 4 we give the full statement for the existence and the uniqueness of solution of the Monge problem and the uniqueness of the solution of the Monge-Kantorovitch problem under the hypothesis that the Wasserstein distance is finite. We have avoided to give the corresponding proofs which are quite technical (cf. [13]), however all the applications are provided with proofs or explanations in such a way that the reader can have an idea about how to do it.

Section 5 studies the Monge-Ampère equation for the measures which are absolutely continuous with respect to the Wiener measure. First we define the Alexandrov versions of the Ornstein-Uhlenbeck operator and the second order Sobolev derivatives for 1-convex Wiener maps. With the help of these, we write the corresponding Jacobian using the modified Carleman-Fredholm determinant which is natural in the infinite dimensional case (cf., [29]). Here we have a major difficulty which originates from the pathology of the Radon-Nikodym derivatives of the vector measures with respect to a scalar measure: in fact even if the second order Sobolev derivative of a Wiener function is a vector measure with values in the space of Hilbert-Schmidt operators, its absolutely continuous part has no reason to be Hilbert-Schmidt. Hence the Carleman-Fredholm determinant may not exist, however due to the 1-convexity, the detereminants of the approximating sequence are all with values in the interval $[0, 1]$. Consequently we can construct the subsolutions with the help of the Fatou lemma.

Last but not the least, in section 6, we remark that all these difficulties can be overcome thanks to the natural renormalization of the Itô stochastic calculus. In fact using the Itô representation theorem and the Wiener space analysis extended to the distributions [27], we can give the explicit solution of the Monge-Ampère equation. This is a remarkable result in the sense that such techniques do not exist in the finite dimensional case.

§2. Preliminaries and notations

Let W be a separable Fréchet space equipped with a Gaussian measure μ of zero mean whose support is the whole space. The corresponding Cameron-Martin space is denoted by H. Recall that the

injection $H \hookrightarrow W$ is compact and its adjoint is the natural injection $W^\star \hookrightarrow H^\star \subset L^2(\mu)$. The triple (W, μ, H) is called an abstract Wiener space. Recall that $W = H$ if and only if W is finite dimensional. A subspace F of H is called regular if the corresponding orthogonal projection has a continuous extension to W, denoted again by the same letter. It is well-known that there exists an increasing sequence of regular subspaces $(F_n, n \geq 1)$, called total, such that $\cup_n F_n$ is dense in H and in W. Let $\sigma(\pi_{F_n})^{3)}$ be the σ-algebra generated by π_{F_n}, then for any $f \in L^p(\mu)$, the martingale sequence $(E[f|\sigma(\pi_{F_n})], n \geq 1)$ converges to f (strongly if $p < \infty$) in $L^p(\mu)$. Observe that the function $f_n = E[f|\sigma(\pi_{F_n})]$ can be identified with a function on the finite dimensional abstract Wiener space (F_n, μ_n, F_n), where $\mu_n = \pi_n \mu$.

Since the translations of μ with the elements of H induce measures equivalent to μ, the Gâteaux derivative in H direction of the random variables is a closable operator on $L^p(\mu)$-spaces and this closure will be denoted by ∇ cf., for example [10, 28]. The corresponding Sobolev spaces (the equivalence classes) of the real random variables will be denoted as $\mathbb{D}_{p,k}$, where $k \in \mathbb{N}$ is the order of differentiability and $p > 1$ is the order of integrability. If the random variables are with values in some separable Hilbert space, say Φ, then we shall define similarly the corresponding Sobolev spaces and they are denoted as $\mathbb{D}_{p,k}(\Phi)$, $p > 1$, $k \in \mathbb{N}$. Since $\nabla : \mathbb{D}_{p,k} \to \mathbb{D}_{p,k-1}(H)$ is a continuous and linear operator its adjoint is a well-defined operator which we represent by δ. In the case of classical Wiener space, i.e., when $W = C(\mathbb{R}_+, \mathbb{R}^d)$, then δ coincides with the Itô integral of the Lebesgue density of the adapted elements of $\mathbb{D}_{p,k}(H)$ (cf. [28]).

For any $t \geq 0$ and measurable $f : W \to \mathbb{R}_+$, we note by

$$P_t f(x) = \int_W f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) \mu(dy),$$

it is well-known that $(P_t, t \in \mathbb{R}_+)$ is a hypercontractive semigroup on $L^p(\mu), p > 1$, which is called the Ornstein-Uhlenbeck semigroup (cf. [10, 28]). Its infinitesimal generator is denoted by $-\mathcal{L}$ and we call $\mathcal{L}$ the Ornstein-Uhlenbeck operator (sometimes called the number operator by the physicists). The norms defined by

$$(2.2) \qquad \|\phi\|_{p,k} = \|(I + \mathcal{L})^{k/2}\phi\|_{L^p(\mu)}$$

are equivalent to the norms defined by the iterates of the Sobolev derivative ∇. This observation permits us to identify the duals of the space

---

3) For the notational simplicity, in the sequel we shall denote it by π_{F_n}.

$\mathbb{D}_{p,k}(\Phi); p > 1, k \in \mathbb{N}$ by $\mathbb{D}_{q,-k}(\Phi')$, with $q^{-1} = 1 - p^{-1}$, where the latter space is defined by replacing k in (2.2) by $-k$, this gives us the distribution spaces on the Wiener space W (in fact we can take as k any real number). An easy calculation shows that, formally, $\delta \circ \nabla = \mathcal{L}$, and this permits us to extend the divergence and the derivative operators to the distributions as linear, continuous operators. In fact $\delta : \mathbb{D}_{q,k}(H \otimes \Phi) \to \mathbb{D}_{q,k-1}(\Phi)$ and $\nabla : \mathbb{D}_{q,k}(\Phi) \to \mathbb{D}_{q,k-1}(H \otimes \Phi)$ continuously, for any $q > 1$ and $k \in \mathbb{R}$, where $H \otimes \Phi$ denotes the completed Hilbert-Schmidt tensor product (cf., for instance [28]).

Let us recall some facts from the convex analysis. Let K be a Hilbert space, a subset S of $K \times K$ is called cyclically monotone if any finite subset $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ of S satisfies the following algebraic condition:

$$\langle y_1, x_2 - x_1 \rangle + \langle y_2, x_3 - x_2 \rangle + \cdots + \langle y_{N-1}, x_N - x_{N-1} \rangle + \langle y_N, x_1 - x_N \rangle \leq 0,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of K. It turns out that S is cyclically monotone if and only if

$$\sum_{i=1}^{N} (y_i, x_{\sigma(i)} - x_i) \leq 0,$$

for any permutation σ of $\{1, \ldots, N\}$ and for any finite subset $\{(x_i, y_i) : i = 1, \ldots, N\}$ of S. Note that S is cyclically monotone if and only if any translate of it is cyclically monotone. By a theorem of Rockafellar, any cyclically monotone set is contained in the graph of the subdifferential of a convex function in the sense of convex analysis ([22]) and even if the function may not be unique its subdifferential is unique.

Let now (W, μ, H) be an abstract Wiener space; a measurable function $f : W \to \mathbb{R} \cup \{\infty\}$ is called 1-convex if the map

$$h \to f(x + h) + \frac{1}{2} |h|_H^2 = F(x, h)$$

is convex on the Cameron-Martin space H with values in $L^0(\mu)$. Note that this notion is compatible with the μ-equivalence classes of random variables thanks to the Cameron-Martin theorem. It is proven in [11] that this definition is equivalent the following condition: Let $(\pi_n, n \geq 1)$ be a sequence of regular, finite dimensional, orthogonal projections of H, increasing to the identity map I_H. Denote also by π_n its continuous extension to W and define $\pi_n^{\perp} = I_W - \pi_n$. For $x \in W$, let $x_n = \pi_n x$ and $x_n^{\perp} = \pi_n^{\perp} x$. Then f is 1-convex if and only if

$$x_n \to \frac{1}{2} |x_n|_H^2 + f(x_n + x_n^{\perp})$$

is $\pi_n^\perp \mu$-almost surely convex.

§3. Some Inequalities

Definition 3.1. *Let ξ and η be two probabilities on $(W, \mathcal{B}(W))$. We say that a probability γ on $(W \times W, \mathcal{B}(W \times W))$ is a solution of the Monge-Kantorovitch problem associated to the couple (ξ, η) if the first marginal of γ is ξ, the second one is η and if*

$$J(\gamma) = \int_{W \times W} |x - y|_H^2 d\gamma(x, y)$$

$$= \inf \left\{ \int_{W \times W} |x - y|_H^2 d\beta(x, y) : \beta \in \Sigma(\xi, \eta) \right\},$$

where $\Sigma(\xi, \eta)$ denotes the set of all the probability measures on $W \times W$ whose first and second marginals are respectively ξ and η. We shall denote the Wasserstein distance between ξ and η, which is the positive square-root of this infimum, with $d_H(\xi, \eta)$.

Remark: By the weak compacteness of probability measures on $W \times W$ and the lower semi-continuity of the strictly convex cost function, the infimum in the definition is attained even if the functional J is identically infinity.

The following result, whose proof is outlined below (cf. also[12, 13]) is an extension of an inequality due to Talagrand [24] and it gives a sufficient condition for the Wasserstein distance to be finite:

Theorem 3.1. *Let $L \in \mathbb{L} \log \mathbb{L}(\mu)$ be a positive random variable with $E[L] = 1$ and let ν be the measure $d\nu = L d\mu$. We then have*

$$(3.3) \qquad d_H^2(\nu, \mu) \le 2E[L \log L].$$

Proof: Let us remark first that we can take W as the classical Wiener space $W = C_0([0, 1])$ and, using the stopping techniques of the martingale theory, we may assume that L is upper and lower bounded almost surely. Then a classical result of the Itô calculus implies that L can be represented as an exponential martingale

$$L_t = \exp \left\{ -\int_0^t \dot{u}_\tau dW_\tau - \frac{1}{2} \int_0^t |\dot{u}_\tau|^2 d\tau \right\},$$

with $L = L_1$. Let us define $u : W \to H$ as $u(t, x) = \int_0^t \dot{u}_\tau d\tau$ and $U : W \to W$ as $U(x) = x + u(x)$. The Girsanov theorem implies that

$x \to U(x)$ is a Browian motion under ν, hence $\beta = (U \times I)\nu \in \Sigma(\mu, \nu)$. Let γ be any optimal measure, then

$$
\begin{aligned}
J(\gamma) &= d_H^2(\nu, \mu) \le \int_{W \times W} |x - y|_H^2 d\beta(x, y) \\
&= E[|u|_H^2 L] \\
&= 2E[L \log L],
\end{aligned}
$$

where the last equality follows also from the Girsanov theorem and the Itô stochastic calculus.   □

Combining Theorem 3.1 with the triangle inequality for the Wasserstein distance gives:

Corollary 3.1. *Assume that ν_i $(i = 1, 2)$ have Radon-Nikodym densities L_i $(i = 1, 2)$ with respect to the Wiener measure μ which are in $\mathbb{L} \log \mathbb{L}$. Then*

$$d_H(\nu_1, \nu_2) < \infty.$$

Let us give a simple application of the above result in the lines of [17]:

Corollary 3.2. *Assume that $A \in \mathcal{B}(W)$ is any set of positive Wiener measure. Define the H-gauge function of A as*

$$q_A(x) = \inf(|h|_H : h \in (A - x) \cap H).$$

Then we have

$$E[q_A^2] \le 2 \log \frac{1}{\mu(A)},$$

in other words

$$\mu(A) \le \exp\left\{ -\frac{E[q_A^2]}{2} \right\}.$$

Similarly if A and B are H-separated, i.e., if $A_\varepsilon \cap B = \emptyset$, for some $\varepsilon > 0$, where $A_\varepsilon = \{x \in W : q_A(x) \le \varepsilon\}$, then

$$\mu(A_\varepsilon^c) \le \frac{1}{\mu(A)} e^{-\varepsilon^2/4}$$

and consequently

$$\mu(A)\,\mu(B) \le \exp\left(-\frac{\varepsilon^2}{4} \right).$$

Remark 3.1. *We already know that, from the $0 - 1$-law, q_A is almost surely finite, besides it satisfies $|q_A(x + h) - q_A(x)| \le |h|_H$, hence $E[\exp \lambda q_A^2] < \infty$ for any $\lambda < 1/2$ (cf. the Appendix B.8 of [29]). In fact all these assertions can also be proved with the technique used below.*

Proof: Let ν_A be the measure defined by

$$d\nu_A = \frac{1}{\mu(A)} 1_A d\mu .$$

Let γ_A be the solution of the Monge-Kantorovitch problem, it is easy to see that the support of γ_A is included in $W \times A$, hence

$$|x - y|_H \geq \inf\{|x - z|_H : z \in A\} = q_A(x) ,$$

γ_A-almost surely. This implies in particular that q_A is almost surely finite. It follows now from the inequality (3.3)

$$E[q_A^2] \leq -2 \log \mu(A) ,$$

hence the proof of the first inequality follows. For the second let $B = A_\varepsilon^c$ and let γ_{AB} be the solution of the Monge-Kantorovitch problem corresponding to ν_A, ν_B. Then we have from the Corollary 3.1,

$$d_H^2(\nu_A, \nu_B) \leq -4 \log \mu(A)\mu(B) .$$

Besides the support of the measure γ_{AB} is in $A \times B$, hence γ_{AB}-almost surely $|x - y|_H \geq \varepsilon$ and the proof follows. $\qquad\square$

§4. Construction of the transport map

In this section we call optimal every probability measure[4] γ on $W \times W$ such that $J(\gamma) < \infty$ and that $J(\gamma) \leq J(\theta)$ for every other probability θ having the same marginals as those of γ. We recall that a finite dimensional subspace F of W is called regular if the corresponding projection is continuous. Similarly a finite dimensional projection of H is called regular if it has a continuous extension to W.

The proof of the next theorem, for which we refer the reader to [13], can be done by choosing a proper disintegration of any optimal measure in such a way that the elements of this disintegration are the solutions of finite dimensional Monge-Kantorovitch problems. The latter is proven with the help of the section-selection theorem [6].

Theorem 4.1 (General case). *Suppose that ρ and ν are two probability measures on W such that*

$$d_H(\rho, \nu) < \infty .$$

---

[4] In fact the results of this section are essentially true for the bounded, positive measures.

Let $(\pi_n, n \geq 1)$ be a total increasing sequence of regular projections (of H, converging to the identity map of H). Suppose that, for any $n \geq 1$, the regular conditional probabilities $\rho(\cdot | \pi_n^\perp = x^\perp)$ vanish $\pi_n^\perp \rho$-almost surely on the subsets of $(\pi_n^\perp)^{-1}(W)$ with Hausdorff dimension $n - 1$. Then there exists a unique solution of the Monge-Kantorovitch problem, denoted by $\gamma \in \Sigma(\rho, \nu)$ and γ is supported by the graph of a Borel map T which is the solution of the Monge problem. $T : W \to W$ is of the form $T = I_W + \xi$, where $\xi \in H$ almost surely. Besides we have

$$d_H^2(\rho, \nu) = \int_{W \times W} |T(x) - x|_H^2 d\gamma(x, y)$$

$$= \int_W |T(x) - x|_H^2 d\rho(x),$$

and for $\pi_n^\perp \rho$-almost almost all $x_n^\perp$, the map $u \to \xi(u + x_n^\perp)$ is cyclically monotone on $(\pi_n^\perp)^{-1}\{x_n^\perp\}$, in the sense that

$$\sum_{i=1}^N \left(u_i + \xi(x_n^\perp + u_i), u_{i+1} - u_i\right)_H \leq 0$$

$\pi_n^\perp \rho$-almost surely, for any cyclic sequence $\{u_1, \ldots, u_N, u_{N+1} = u_1\}$ from $\pi_n(W)$. Finally, if, for any $n \geq 1$, $\pi_n^\perp \nu$-almost surely, $\nu(\cdot | \pi_n^\perp = y^\perp)$ also vanishes on the $n-1$-Hausdorff dimensional subsets of $(\pi_n^\perp)^{-1}(W)$, then T is invertible, i.e, there exists $S : W \to W$ of the form $S = I_W + \eta$ such that $\eta \in H$ satisfies a similar cyclic monotononicity property as ξ and that

$$1 = \gamma\{(x, y) \in W \times W : T \circ S(y) = y\}$$

$$= \gamma\{(x, y) \in W \times W : S \circ T(x) = x\}.$$

In particular we have

$$d_H^2(\rho, \nu) = \int_{W \times W} |S(y) - y|_H^2 d\gamma(x, y)$$

$$= \int_W |S(y) - y|_H^2 d\nu(y).$$

Remark 4.1. *In particular, for all the measures ρ which are absolutely continuous with respect to the Wiener measure μ, the second hypothesis is satisfied, i.e., the measure $\rho(\cdot | \pi_n^\perp = x_n^\perp)$ vanishes on the sets of Hausdorff dimension $n - 1$.*

The case where one of the measures is the Wiener measure and the other is absolutely continuous with respect to μ is the most important

one for the applications. Consequently we give the related results separately in the following theorem where the tools of the Malliavin calculus give more information about the maps ξ and η of Theorem 4.1:

Theorem 4.2 (Gaussian case). *Let ν be the measure $d\nu = Ld\mu$, where L is a positive random variable, with $E[L] = 1$. Assume that $d_H(\mu, \nu) < \infty$ (for instance $L \in \mathbb{L} \log \mathbb{L}$). Then there exists a 1-convex function $\phi \in \mathbb{D}_{2,1}$, unique upto a constant, such that the map $T = I_W + \nabla\phi$ is the unique solution of the original problem of Monge. Moreover, its graph supports the unique solution of the Monge-Kantorovitch problem γ. Consequently*

$$(I_W \times T)\mu = \gamma$$

In particular T maps μ to ν and T is almost surely invertible, i.e., there exists some T^{-1} such that $T^{-1}\nu = \mu$ and that

$$
\begin{aligned}
1 &= \mu\left\{x : T^{-1} \circ T(x) = x\right\} \\
&= \nu\left\{y \in W : T \circ T^{-1}(y) = y\right\}.
\end{aligned}
$$

Remark 4.2. *Assume that the operator ∇ is closable with respect to ν, then we have $\eta = \nabla\psi$. In particular, if ν and μ are equivalent, then we have*

$$T^{-1} = I_W + \nabla\psi,$$

where is ψ is a 1-convex function.

Remark 4.3. *Let $(e_n, n \in \mathbb{N})$ be a complete, orthonormal in H, denote by V_n the sigma algebra generated by $\{\delta e_1, \ldots, \delta e_n\}$ and let $L_n = E[L|V_n]$. If $\phi_n \in \mathbb{D}_{2,1}$ is the function constructed in Theorem 4.2, corresponding to L_n, then, using the inequality (3.3) we can prove that the sequence $(\phi_n, n \in \mathbb{N})$ converges to ϕ in $\mathbb{D}_{2,1}$.*

Remark 4.4. *Assume that $L \in \mathbb{L}^1_+(\mu)$, with $E[L] = 1$ and let $(D_k, k \in \mathbb{N})$ be a measurable partition of W such that on each D_k, L is bounded. Define $d\nu = L\,d\mu$ and $\nu_k = \nu(\cdot|D_k)$. It follows from Theorem 3.1, that $d_H(\mu, \nu_k) < \infty$. Let then T_k be the map constructed in Theorem 4.2 satisfying $T_k\mu = \nu_k$. Define $n(dk)$ as the probability distribution on $\mathbb{N}$ given by $n(\{k\}) = \nu(D_k)$, $k \in \mathbb{N}$. Then we have*

$$\int_W f(y)d\nu(y) = \int_{W\times\mathbb{N}} f(T_k(x))\mu(dx)n(dk).$$

A similar result is given in [9], the difference with that of above lies in the fact that we have a more precise information about the probability space on which T is defined.

Let us give some applications of the above theorem to the factorization of the absolutely continuous transformations of the Wiener measure.

Assume that $V = I_W + v : W \to W$ be an absolutely continuous transformation and let $L \in \mathbb{L}^1_+(\mu)$ be the Radon-Nikodym derivative of $V\mu$ with respect to μ. Let $T = I_W + \nabla\phi$ be the transport map such that $T\mu = L.\mu$. Then it is easy to see that the map $s = T^{-1} \circ V$ is a rotation, i.e., $s\mu = \mu$ (cf. [29]) and it can be represented as $s = I_W + \alpha$. In particular we have

$$(4.4) \qquad\qquad \alpha + \nabla\phi \circ s = v.$$

Since ϕ is a 1-convex map, we have $h \to \frac{1}{2}|h|_H^2 + \phi(x+h)$ is almost surely convex (cf. [11]). Let $s' = I_W + \alpha'$ be another rotation with $\alpha' : W \to H$. By the 1-convexity of ϕ, we have

$$\frac{1}{2}|\alpha'|_H^2 + \phi \circ s' \geq \frac{1}{2}|\alpha|_H^2 + \phi \circ s + \left(\alpha + \nabla\phi \circ s, \alpha' - \alpha\right)_H,$$

μ-almost surely. Taking the expectation of both sides, using the fact that s and s' preserve the Wiener measure μ and the identity (4.4), we obtain

$$E\left[\frac{1}{2}|\alpha|_H^2 - (v, \alpha)_H\right] \leq E\left[\frac{1}{2}|\alpha'|_H^2 - (v, \alpha')_H\right].$$

Hence we have proven the existence part of the following

Proposition 4.1. *Let $\mathcal{R}_2$ denote the subset of $L^2(\mu, H)$ whose elements are defined by the property that $x \to x + \eta(x)$ is a rotation, i.e., it preserves the Wiener measure. Then α is the unique element of $\mathcal{R}_2$ which minimizes the functional*

$$\eta \to M_v(\eta) = E\left[\frac{1}{2}|\eta|_H^2 - (v, \eta)_H\right].$$

Proof: The only claim to prove is the uniqueness and it follows easily from Theorem 4.2. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The following theorem, whose proof is rather easy, gives a better understanding of the structure of absolutely continuous transformations of the Wiener measure:

Theorem 4.3. *Assume that $U : W \to W$ be a measurable map and $L \in \mathbb{L}\log\mathbb{L}$ a positive random variable with $E[L] = 1$. Assume that the measure $\nu = L \cdot \mu$ is a Girsanov measure for U, i.e., that one has*

$$E[f \circ U\, L] = E[f],$$

for any $f \in C_b(W)$. Then there exists a unique map $T = I_W + \nabla\phi$ with $\phi \in \mathbb{D}_{2,1}$ is 1-convex, and a measure preserving transformation $R : W \to W$ such that $U \circ T = R$ μ-almost surely and $U = R \circ T^{-1}$ ν-almost surely.

Another version of Theorem 4.3 can be announced as follows:

Theorem 4.4. *Assume that $Z : W \to W$ is a measurable map such that $Z\mu \ll \mu$, with $d_H(Z\mu, \mu) < \infty$. Then Z can be decomposed as*

$$Z = T \circ s\,,$$

where T is the unique transport map of the Monge-Kantorovitch problem for $\Sigma(\mu, Z\mu)$ and s is a rotation.

Although the following result is a translation of the results of this section, it is interesting from the point of view of stochastic differential equations:

Theorem 4.5. *Let (W, μ, H) be the standard Wiener space on $\mathbb{R}^d$, i.e., $W = C(\mathbb{R}_+, \mathbb{R}^d)$. Assume that there exists a probability $P \ll \mu$ which is the weak solution of the stochastic differential equation*

$$dy_t = dW_t + b(t, y)dt\,,$$

such that $d_H(P, \mu) < \infty$. Then there exists a process $(T_t, t \in \mathbb{R}_+)$ which is a pathwise solution of some (anticipative) stochastic differential equation whose law is equal to P.

Proof: Let T be the transport map constructed in Theorem 4.2 corresponding to $dP/d\mu$. Then it has an inverse T^{-1} such that $\mu\{T^{-1} \circ T(x) = x\} = 1$. Let ϕ be the 1-convex function such that $T = I_W + \nabla\phi$ and denote by $(D_s\phi, s \in \mathbb{R}_+)$ the representation of $\nabla\phi$ in $L^2(\mathbb{R}_+, ds)$. Define $T_t(x)$ as the trajectory $T(x)$ evaluated at $t \in \mathbb{R}_+$. Then it is easy to see that $(T_t, t \in \mathbb{R}_+)$ satisfies the stochastic differential equation

$$T_t(x) = W_t(x) + \int_0^t l(s, T(x))ds\,,\ t \in \mathbb{R}_+\,,$$

where $W_t(x) = x(t)$ and $l(s, x) = D_s\phi \circ T^{-1}(x)$ if $x \in T(W)$ and zero otherwise. $\qquad\square$

§5.   The Monge-Ampère equation

Assume that $W = \mathbb{R}^n$ and take a density $L \in \mathbb{L} \log \mathbb{L}$. Let $\phi \in \mathbb{D}_{2,1}$ be the 1-convex function such that $T = I + \nabla\phi$ maps μ to $L \cdot \mu$. Let $S = I + \nabla\psi$ be its inverse with $\psi \in \mathbb{D}_{2,1}$. Let now $\nabla_a^2\phi$ be the second Alexandrov derivative of ϕ, i.e., the Radon-Nikodym derivative of the absolutely continuous part of the vector measure $\nabla^2\phi$ with respect to the Gaussian measure μ on $\mathbb{R}^n$. Since ϕ is 1-convex, it follows that $\nabla^2\phi \geq -I_{\mathbb{R}^n}$ in the sense of the distributions, consequently $\nabla_a^2\phi \geq -I_{\mathbb{R}^n}$ μ-almost surely. Define also the Alexandrov version $\mathcal{L}_a\phi$ of $\mathcal{L}\phi$ as the Radon-Nikodym derivative of the absolutely continuous part of the distribution $\mathcal{L}\phi$. Since we are in finite dimensional situation, we have the explicit expression for $\mathcal{L}_a\phi$ as

$$\mathcal{L}_a\phi(x) = (\nabla\phi(x), x)_{\mathbb{R}^n} - \text{trace}\left(\nabla_a^2\phi\right).$$

Let Λ be the Gaussian Jacobian

$$\Lambda = \det{}_2\left(I_{\mathbb{R}^n} + \nabla_a^2\phi\right)\exp\left\{-\mathcal{L}_a\phi - \frac{1}{2}|\nabla\phi|_{\mathbb{R}^n}^2\right\}.$$

Remark 5.1. In this expression as well as in the sequel, the notation $\det_2(I_H + A)$ denotes the modified Carleman-Fredholm determinant of the operator $I_H + A$ on a Hilbert space H. If A is an operator of finite rank, then it is defined as

$$\det{}_2\left(I_H + A\right) = \prod_{i=1}^{n}(1 + l_i)e^{-l_i},$$

where $(l_i, i \leq n)$ denotes the eigenvalues of A counted with respect to their multiplicity. In fact this determinant has an analytic extension to the space of Hilbert-Schmidt operators on a separable Hilbert space, cf. [7] and Appendix A.2 of [29]. As explained in [29], the modified determinant exists for the Hilbert-Schmidt operators while the ordinary determinant does not, since the latter requires the existence of the trace of A. Hence the modified Carleman-Fredholm determinant is particularly useful when one studies the absolute continuity properties of the image of a Gaussian measure under non-linear transformations in the setting of infinite dimensional Banach spaces (cf., [29] for further information).

It follows from the change of variables formula given in Corollary 4.3 of [19], that, for any $f \in C_b(\mathbb{R}^n)$,

$$E[f \circ T \Lambda] = E\left[f\, 1_{\partial\Phi(M)}\right],$$

where M is the set of non-degeneracy of $I_{\mathbb{R}^n} + \nabla_a^2 \phi$,

$$\Phi(x) = \frac{1}{2}|x|^2 + \phi(x)$$

and $\partial \Phi$ denotes the subdifferential of the convex function Φ. Let us note that, in case $L > 0$ almost surely, T has a global inverse S, i.e., $S \circ T = T \circ S = I_{\mathbb{R}^n}$ μ-almost surely and $\mu(\partial\Phi(M)) = \mu(S^{-1}(M))$. Assume now that $\Lambda > 0$ almost surely, i.e., that $\mu(M) = 1$. Then, for any $f \in C_b(\mathbb{R}^n)$, we have

$$
\begin{aligned}
E[f \circ T] &= E\left[f \circ T \frac{\Lambda}{\Lambda \circ T^{-1} \circ T}\right] \\
&= E\left[f \frac{1}{\Lambda \circ T^{-1}} 1_{\partial\Phi(M)}\right] \\
&= E[f\, L],
\end{aligned}
$$

where T^{-1} denotes the left inverse of T whose existence is guaranteed by Theorem 4.2. Since $T(x) \in \partial\Phi(M)$ almost surely, it follows from the above calculations

$$\frac{1}{\Lambda} = L \circ T,$$

almost surely. Take now any $t \in [0, 1)$, the map $x \to \frac{1}{2}|x|_H^2 + t\phi(x) = \Phi_t(x)$ is strictly convex and a simple calculation implies that the mapping $T_t = I + t\nabla\phi$ is $(1-t)$-monotone (cf. [29], Chapter 6), consequently it has a left inverse denoted by S_t. Let us denote by Ψ_t the Legendre transformation of Φ_t:

$$\Psi_t(y) = \sup_{x \in \mathbb{R}^n} \{(x, y) - \Phi_t(x)\}.$$

A simple calculation shows that

$$
\begin{aligned}
\Psi_t(y) &= \sup_x \left[(1-t)\left\{(x, y) - \frac{|x|^2}{2}\right\} + t\left\{(x, y) - \frac{|x|^2}{2} - \phi(x)\right\}\right] \\
&\leq (1-t)\frac{|y|^2}{2} + t\Psi_1(y).
\end{aligned}
$$

Since Ψ_1 is the Legendre transformation of $\Phi_1(x) = |x|^2/2 + \phi(x)$ and since $L \in \mathbb{L}\log\mathbb{L}$, it is finite on a convex set of full measure, hence it is finite everywhere. Consequently $\Psi_t(y) < \infty$ for any $y \in \mathbb{R}^n$. Since a finite, convex function is almost everywhere differentiable, $\nabla\Psi_t$ exists almost everywhere on and it is equal almost everywhere on $T_t(M_t)$ to the left inverse T_t^{-1}, where M_t is the set of non-degeneracy of $I_{\mathbb{R}^n} + t\nabla_a^2\phi$.

Note that $\mu(M_t) = 1$. The strict convexity implies that T_t^{-1} is Lipschitz with a Lipschitz constant $\frac{1}{1-t}$. Let now Λ_t be the Gaussian Jacobian

$$\Lambda_t = \det{}_2\left(I_{\mathbb{R}^n} + t\nabla_a^2\phi\right)\exp\left\{-t\mathcal{L}_a\phi - \frac{t^2}{2}|\nabla\phi|_{\mathbb{R}^n}^2\right\}.$$

Since the domain of ϕ is the whole space $\mathbb{R}^n$, $\Lambda_t > 0$ almost surely, hence, as we have explained above, it follows from the change of variables formula of [19] that $T_t\mu$ is absolutely continuous with respect to μ and that

$$\frac{1}{\Lambda_t} = L_t \circ T_t\,,$$

μ-almost surely.

Let us come back to the infinite dimensional case: we first give an inequality which may be useful.

Theorem 5.1. *Assume that (W, μ, H) is an abstract Wiener space, assume that $K, L \in \mathbb{L}_+^1(\mu)$ with $K > 0$ almost surely and denote by $T : W \to W$ the transfer map $T = I_W + \nabla\phi$, which maps the measure $Kd\mu$ to the measure $Ld\mu$. Then the following inequality holds:*

$$(5.5) \qquad \frac{1}{2}E[|\nabla\phi|_H^2] \le E[-\log K + \log L \circ T]\,.$$

Proof: Let us define k as $k = K \circ T^{-1}$, then for any $f \in C_b(W)$, we have

$$\int_W f(y)L(y)d\mu(y) = \int_W f \circ T(x)K(x)d\mu(x)$$
$$= \int_W f \circ T(x)k \circ T(x)d\mu(x)\,,$$

hence

$$T\mu = \frac{L}{k}\cdot\mu\,.$$

It then follows from the inequality (3.3) that

$$\frac{1}{2}E\left[|\nabla\phi|_H^2\right] \le E\left[\frac{L}{k}\log\frac{L}{k}\right]$$
$$= E\left[\log\frac{L \circ T}{k \circ T}\right]$$
$$= E[-\log K + \log L \circ T]\,.$$

$\square$

In case K and L are given as e^{-U} and e^{-V} we have another inequality containing the Fisher information (cf. [21] for the finite dimensional case):

Theorem 5.2. *Assume that $U, V \in \mathbb{D}_{2,1}$ are such that $E[e^{-U}] = E[e^{-V}] = 1$ and that $E[e^{-U}|\nabla U|_H^2] + E[e^{-V}|\nabla V|_H^2] < \infty$. Define $d\rho = e^{-U}d\mu$ and $d\nu = e^{-V}d\mu$. Then we have*

$$(5.6) \qquad d_H(\rho, \nu) \leq E_\rho[|\nabla U|_H^2]^{1/2} + E_\nu[|\nabla V|_H^2]^{1/2}.$$

The equation (5.6) can be refined in the following way: let

$$\kappa_\pm(U,V) = E_\rho[|\nabla U|_H^2]^{1/2} \pm \left\{ E_\rho[|\nabla U|_H^2] + 2(E_\rho(U) - E_\nu(V)) \right\}^{1/2}$$

and let

$$\kappa_\pm(V,U) = E_\nu[|\nabla V|_H^2]^{1/2} \pm \left\{ E_\nu[|\nabla V|_H^2] + 2(E_\nu(V) - E_\rho(U)) \right\}^{1/2}.$$

We then have
(5.7)
$$\max\{\kappa_-(U,V), \kappa_-(V,U)\} \leq d_H(\rho,\nu) \leq \min\{\kappa_+(U,V), \kappa_+(V,U)\}.$$

Proof: By taking the conditional expectations with respect to the sigma algebras generated by a finite number of elements of the first Wiener chaos and using the Jensen inequality, we can reduce the problem to the finite dimensional case. Let now $T = I + \nabla\phi$ be the transport map sending ρ to ν, let also $S = I + \nabla\psi$ be its inverse. As we have seen before, we can write $\Lambda = \exp(-U + V \circ T)$ as

$$\Lambda = \det_2(I + \nabla_a^2\phi) \exp\left\{ -\mathcal{L}_a\phi - \frac{1}{2}|\nabla\phi|_H^2 \right\}.$$

Solving $|\nabla\phi|_H^2$ from this expression and using the fact that $-\mathcal{L}_a\phi \leq -\mathcal{L}\phi$ in the sense of the distributions, we obtain

$$\frac{1}{2}|\nabla\phi|_H^2 \leq U - V \circ T - \mathcal{L}\phi.$$

Taking the expectation of both sides of this inequality with respect to ρ, we obtain

$$(5.8) \qquad \frac{1}{2}d_H^2(\rho,\nu) \leq E_\rho[U] - E_\nu[V] + E_\rho[(\nabla\phi, \nabla U)].$$

Interchanging T and S and ρ and ν in the inequality (5.8), we obtain also

$$(5.9) \qquad \frac{1}{2}d_H^2(\rho,\nu) \leq E_\nu[V] - E_\rho[U] + E_\nu[(\nabla\psi, \nabla V)].$$

Adding (5.8) to (5.9) and using the Cauchy-Schwarz inequality completes the proof of the first part. For the second, using the inequalities (5.8) and (5.9), with the help of the Cauchy-Schwarz inequality and the general expression of solutions of the second order polynomial equation we get the claim at once. □

Suppose that $\phi \in \mathbb{D}_{2,1}$ is a 1-convex Wiener functional. Let V_n be, as usual, the sigma algebra generated by $\{\delta e_1, \ldots, \delta e_n\}$, where $(e_n, n \geq 1)$ is an orthonormal basis of the Cameron-Martin space H. Then $\phi_n = E[\phi|V_n]$ is again 1-convex (cf. [11]), hence $\mathcal{L}\phi_n$ is a measure as it can be easily verified. However the sequence $(\mathcal{L}\phi_n, n \geq 1)$ converges to $\mathcal{L}\phi$ only in $\mathbb{D}'$. Consequently, there is no reason for the limit $\mathcal{L}\phi$ to be a measure. In case this happens, we shall denote the Radon-Nikodym density with respect to μ, of the absolutely continuous part of this measure by $\mathcal{L}_a\phi$.

Lemma 5.1. *Let $\phi \in \mathbb{D}_{2,1}$ be 1-convex and let V_n be defined as above and define $F_n = E[\phi|V_n]$. Then the sequence $(\mathcal{L}_a F_n, n \geq 1)$ is a submartingale, where $\mathcal{L}_a F_n$ denotes the μ-absolutely continuous part of the measure $\mathcal{L}F_n$.*

Proof: Note that, due to the 1-convexity, we have $\mathcal{L}_a F_n \geq \mathcal{L}F_n$ for any $n \in \mathbb{N}$. Let $X_n = \mathcal{L}_a F_n$ and $f \in \mathbb{D}$ be a positive, V_n-measurable test function. Since $\mathcal{L}E[\phi|V_n] = E[\mathcal{L}\phi|V_n]$, we have

$$
\begin{aligned}
E[X_{n+1} f] &\geq \langle \mathcal{L}F_{n+1}, f \rangle \\
&= \langle \mathcal{L}F_n, f \rangle,
\end{aligned}
$$

where $\langle \cdot, \cdot \rangle$ denotes the duality bracket for the dual pair $(\mathbb{D}', \mathbb{D})$. Consequently

$$
E[f\, E[X_{n+1}|V_n]] \geq \langle \mathcal{L}F_n, f \rangle,
$$

for any positive, V_n-measurable test function f, it follows that the absolutely continuous part of $\mathcal{L}F_n$ is also dominated by the same conditional expectation and this proves the submartingale property. □

Lemma 5.2. *Assume that $L \in \mathbb{L} \log \mathbb{L}$ is a positive random variable whose expectation is one. Assume further that it is lower bounded by a constant $a > 0$. Let $T = I_W + \nabla\phi$ be the transport map such that $T\mu = L.\mu$ and let $T^{-1} = I_W + \nabla\psi$. Then $\mathcal{L}\psi$ is a Radon measure on $(W, \mathcal{B}(W))$. If L is upper bounded by $b > 0$, then $\mathcal{L}\phi$ is also a Radon measure on $(W, \mathcal{B}(W))$.*

Proof: Let $L_n = E[L|V_n]$, then $L_n \geq a$ almost surely. Let $T_n = I_W + \nabla\phi_n$ be the transport map which satisfies $T_n\mu = L_n.\mu$ and let

$T_n^{-1} = I_W + \nabla \psi_n$ be its inverse. We have

$$L_n = \det_2 \left(I_H + \nabla_a^2 \psi_n \right) \exp \left[-\mathcal{L}_a \psi_n - \frac{1}{2} |\nabla \psi_n|_H^2 \right] .$$

By the hypothesis $- \log L_n \leq - \log a$. Since ψ_n is 1-convex, it follows from the finite dimensional results that $\det_2 \left(I_H + \nabla_a^2 \psi_n \right) \in [0, 1]$ almost surely. Therefore we have

$$\mathcal{L}_a \psi_n \leq - \log a ,$$

besides $\mathcal{L} \psi_n \leq \mathcal{L}_a \psi_n$ as distributions, consequently

$$\mathcal{L} \psi_n \leq - \log a$$

as distributions, for any $n \geq 1$. Since $\lim_n \mathcal{L} \psi_n = \mathcal{L} \psi$ in $\mathbb{D}'$, we obtain $\mathcal{L} \psi \leq - \log a$, hence $- \log a - \mathcal{L} \psi \geq 0$ as a distribution, hence $\mathcal{L} \psi$ is a Radon measure on W, c.f., [10], [28]. This proves the first claim. Note that whenever L is upperbounded, $\Lambda = 1/L \circ T$ is lowerbounded, hence the proof of the second claim is similar to that of the first one. $\square$

Theorem 5.3. *Assume that L is a strictly positive bounded random variable with $E[L] = 1$. Let $\phi \in \mathbb{D}_{2,1}$ be the 1-convex Wiener functional such that*

$$T = I_W + \nabla \phi$$

is the transport map realizing the measure $L . \mu$ and let $S = I_W + \nabla \psi$ be its inverse. Define $F_n = E[\phi | V_n]$, then the submartingale $(\mathcal{L}_a F_n, n \geq 1)$ converges almost surely to $\mathcal{L}_a \phi$. Let $\lambda(\phi)$ be the random variable defined as

$$
\begin{aligned}
\lambda(\phi) &= \liminf_{n \to \infty} \Lambda_n \\
&= \left(\liminf_n \det_2 \left(I_H + \nabla_a^2 F_n \right) \right) \exp \left\{ -\mathcal{L}_a \phi - \frac{1}{2} |\nabla \phi|_H^2 \right\}
\end{aligned}
$$

where

$$\Lambda_n = \det_2 \left(I_H + \nabla_a^2 F_n \right) \exp \left\{ -\mathcal{L}_a F_n - \frac{1}{2} |\nabla F_n|_H^2 \right\} .$$

Then it holds true that

(5.10) $$E[f \circ T \, \lambda(\phi)] \leq E[f]$$

for any $f \in C_b^+(W)$, in particular $\lambda(\phi) \leq \frac{1}{L \circ T}$ almost surely. If $E[\lambda(\phi)] = 1$, then the inequality in (5.10) becomes an equality and we also have

$$\lambda(\phi) = \frac{1}{L \circ T}.$$

Proof: Let us remark that, due to the 1-convexity, $0 \leq \det_2\left(I_H + \nabla_a^2 F_n\right) \leq 1$, hence the lim inf exists. Now, Lemma 5.2 implies that $\mathcal{L}\phi$ is a Radon measure. Let $F_n = E[\phi|V_n]$, then we know from Lemma 5.1 that $(\mathcal{L}_a F_n, n \geq 1)$ is a submartingale. Let $\mathcal{L}^+\phi$ denote the positive part of the measure $\mathcal{L}\phi$. Since $\mathcal{L}^+\phi \geq \mathcal{L}\phi$, we have also $E[\mathcal{L}^+\phi|V_n] \geq E[\mathcal{L}\phi|V_n] = \mathcal{L}F_n$. This implies that $E[\mathcal{L}^+\phi|V_n] \geq \mathcal{L}_a^+ F_n$. Hence we find that

$$\sup_n E[\mathcal{L}_a^+ F_n] < \infty$$

and this condition implies that the submartingale $(\mathcal{L}_a F_n, n \geq 1)$ converges almost surely. We shall now identify the limit of this submartingale. Let $\mathcal{L}_s G$ be the singular part of the measure $\mathcal{L}G$ for a Wiener function G such that $\mathcal{L}G$ is a measure. We have

$$
\begin{aligned}
E[\mathcal{L}\phi|V_n] &= E[\mathcal{L}_a\phi|V_n] + E[\mathcal{L}_s\phi|V_n] \\
&= \mathcal{L}_a F_n + \mathcal{L}_s F_n,
\end{aligned}
$$

hence

$$\mathcal{L}_a F_n = E[\mathcal{L}_a\phi|V_n] + E[\mathcal{L}_s\phi|V_n]_a$$

almost surely, where $E[\mathcal{L}_s\phi|V_n]_a$ denotes the absolutely continuous part of the measure $E[\mathcal{L}_s\phi|V_n]$. Note that, from the Theorem of Jessen (cf., for example Theorem 1.2.1 of [29]), $\lim_n E[\mathcal{L}_s^+\phi|V_n]_a = 0$ and $\lim_n E[\mathcal{L}_s^-\phi|V_n]_a = 0$ almost surely, hence we have

$$\lim_n \mathcal{L}_a F_n = \mathcal{L}_a\phi,$$

μ-almost surely. To complete the proof, an application of the Fatou lemma implies that

$$
\begin{aligned}
E[f \circ T \, \lambda(\phi)] &\leq E[f] \\
&= E\left[f \circ T \frac{1}{L \circ T}\right],
\end{aligned}
$$

for any $f \in C_b^+(W)$. Since T is invertible, it follows that

$$\lambda(\phi) \leq \frac{1}{L \circ T}$$

almost surely. Therefore, in case $E[\lambda(\phi)] = 1$, we have

$$\lambda(\phi) = \frac{1}{L \circ T},$$

and this completes the proof. $\qquad\square$

Corollary 5.1. *Assume that K, L are two positive random variables with values in a bounded interval $[a, b] \subset (0, \infty)$ such that $E[K] = E[L] = 1$. Let $T = I_W + \nabla\phi$, $\phi \in \mathbb{D}_{2,1}$, be the transport map pushing $Kd\mu$ to $Ld\mu$, i.e, $T(Kd\mu) = Ld\mu$. We then have*

$$L \circ T \, \lambda(\phi) \leq K \, ,$$

μ-almost surely. In particular, if $E[\lambda(\phi)] = 1$, then T is the solution of the Monge-Ampère equation.

Proof: Since $a > 0$,

$$\frac{dT\mu}{d\mu} = \frac{L}{K \circ T} \leq \frac{b}{a} \, .$$

Hence, Theorem 5.10 implies that

$$
\begin{aligned}
E[f \circ T \, L \circ T \, \lambda(\phi)] &\leq E[f \, L] \\
&= E[f \circ T \, K] \, ,
\end{aligned}
$$

consequently

$$L \circ T \, \lambda(\phi) \leq K \, ,$$

the rest of the claim is now obvious. $\qquad\square$

For later use we give also the folowing result:

Theorem 5.4. *Assume that L is a positive random variable of class $\mathbb{L} \log \mathbb{L}$ such that $E[L] = 1$. Let $\phi \in \mathbb{D}_{2,1}$ be the 1-convex function corresponding to the transport map $T = I_W + \nabla\phi$. Define $T_t = I_W + t\nabla\phi$, where $t \in [0, 1]$. Then, for any $t \in [0, 1]$, $T_t\mu$ is absolutely continuous with respect to the Wiener measure μ.*

Proof: Let ϕ_n be defined as the transport map corresponding to $L_n = E[P_{1/n}L_n | V_n]$ and define T_n as $I_W + \nabla\phi_n$. For $t \in [0, 1)$, let $T_{n,t} = I_W + t\nabla\phi_n$. It follows from the finite dimensional results which are summarized in the beginning of this section, that $T_{n,t}\mu$ is absolutely

continuous with respect to μ. Let $L_{n,t}$ be the corresponding Radon-Nikodym density and define $\Lambda_{n,t}$ as

$$\Lambda_{n,t} = \det{}_2\left(I_H + t\nabla_a^2\phi_n\right)\exp\left\{-t\mathcal{L}_a\phi_n - \frac{t^2}{2}|\nabla\phi_n|_H^2\right\}.$$

Besides, for any $t \in [0,1)$,

(5.11) $\left((I_H + t\nabla_a^2\phi_n)h, h\right)_H > 0,$

μ-almost surely for any $0 \neq h \in H$. Since ϕ_n is of finite rank, (5.11) implies that $\Lambda_{n,t} > 0$ μ-almost surely and we have shown at the beginning of this section

$$\Lambda_{n,t} = \frac{1}{L_{n,t} \circ T_{n,t}}$$

μ-almost surely. An easy calculation shows that $t \to \Lambda_{n,t}$ is logarithmically concave. Consequently

$$
\begin{aligned}
E\left[L_{t,n}\log L_{t,n}\right] &= E\left[\log L_{n,t} \circ T_{n,t}\right] \\
&= -E\left[\log \Lambda_{t,n}\right] \\
&\leq E\left[L_n \log L_n\right] \\
&\leq E[L\log L],
\end{aligned}
$$

by the Jensen inequality. Therefore

$$\sup_n E[L_{n,t}\log L_{n,t}] < \infty$$

and this implies that the sequence $(L_{n,t}, n \geq 1)$ is uniformly integrable for any $t \in [0,1]$. Consequently it has a subsequence which converges weakly in $L^1(\mu)$ to some L_t. Since, from Theorem 4.2 and from Remark 4, $\lim_n \phi_n = \phi$ in $\mathbb{D}_{2,1}$, where ϕ is the transport map associated to L, for any $f \in C_b(W)$, we have

$$
\begin{aligned}
E[f \circ T_t] &= \lim_k E\left[f \circ T_{n_k,t}\right] \\
&= \lim_k E\left[f\, L_{n_k,t}\right] \\
&= E[f\, L_t],
\end{aligned}
$$

and this completes the proof. $\square$

Let us give an application of this result:

Proposition 5.1. *Assume that the hypothesis of Theorem 5.4 are valid. Let $\nu_t = T_t \mu$ with $\nu_1 = \nu$, $t \in [0,1]$. Then*

$$d_H(\nu_s, \nu_t) = |t - s| d_H(\mu, \nu) \text{ for } s, t \in [0,1].$$

In particular $T_t \circ S_s$ is the optimal transport of ν_s to ν_t and the Wiener functional $\nabla \psi_s + t \nabla \phi \circ S_s$ is an exact form.

Proof: It suffices to prove the claim for the case $s = 1$. Let $\nu_t = T_t \circ S\nu$, then

$$
\begin{aligned}
d_H^2(\nu_t, \nu) &\leq \int |T_t \circ S(y) - y|_H^2 d\nu(y) \\
&= \int |T_t(x) - T(x)|_H^2 d\mu(x) \\
&= (1-t)^2 E[|\nabla \phi|_H^2],
\end{aligned}
$$

which means that

$$d_H(\nu_t, \nu) \leq (1-t) d_H(\mu, \nu).$$

Moreover, from the triangle inequality

$$
\begin{aligned}
d_H(\mu, \nu) &\leq d_H(\nu, \nu_t) + d_H(\nu_t, \mu) \\
&= d_H(\nu, \nu_t) + t d_H(\mu, \nu),
\end{aligned}
$$

we obtain that $(1-t) d_H(\mu, \nu) \leq d_H(\nu, \nu_t)$. The rest is obvious from Theorem 4.2. $\qquad\square$

§6. Solution of the Monge-Ampère equation with the Itô calculus

To have a better understanding of what will follow, let us give an interpretation of the Monge-Ampère equation. Assume that we are given two probability densities K and L and we look for a map $T : W \to W$ such that

$$L \circ T \, J(T) = K$$

almost surely, where $J(T)$ is a kind of Jacobian to be written in terms of T. In Corollary 5.1, we have shown the existence of some $\lambda(\phi)$ which gives an inequality instead of the equality. The reason for that originates from the singularity of the second derivative of the potential function ϕ. Although in the finite dimensional case there are some regularity results about the transport map (cf., [5]), in the infinite dimensional case such

techniques do not work. All these difficulties can be circumvented using the miraculous renormalization of the Itô calculus. In fact assume that K and L satisfy the hypothesis of the corollary. First let us indicate that we can assume $W = C_0([0,1], \mathbb{R})$ (cf., [29], Chapter II, to see how one can pass from an abstract Wiener space to the standard one) and in this case the Cameron-Martin space H becomes $H^1([0,1])$, which is the space of absolutely continuous functions on $[0,1]$, with a square integrable Sobolev derivative. Let now

$$\Lambda = \frac{K}{L \circ T},$$

where T is as constructed above. Then $\Lambda.\mu$ is a Girsanov measure for the map T. This means that the law of the stochastic process $(t,x) \to T_t(x)$ under $\Lambda.\mu$ is equal to the Wiener measure, where $T_t(x)$ is defined as the evaluation of the trajectory $T(x)$ at $t \in [0,1]$. In other words the process $(t,x) \to T_t(x)$ is a Brownian motion under the probability $\Lambda.\mu$. Let $(\mathcal{F}_t^T, t \in [0,1])$ be its filtration, the invertibility of T implies that

$$\bigvee_{t \in [0,1]} \mathcal{F}_t^T = \mathcal{B}(W).$$

Λ is upper and lower bounded μ-almost surely, hence also $\Lambda.\mu$-almost surely. The Itô representation theorem implies that it can be represented as

$$\Lambda = E[\Lambda^2] \exp\left\{ -\int_0^1 \dot{\alpha}_s dT_s - \frac{1}{2} \int_0^1 |\dot{\alpha}_s|^2 ds \right\},$$

where $\alpha(\cdot) = \int_0^\cdot \dot{\alpha}_s ds$ is an H-valued random variable. In fact α can be calculated explicitly using the Itô-Clark representation theorem (cf., [28]), and it is given as

(6.12) $$\dot{\alpha}_t = \frac{E_\Lambda[D_t \Lambda | \mathcal{F}_t^T]}{E_\Lambda[\Lambda | \mathcal{F}_t^T]}$$

$dt \times \Lambda d\mu$-almost surely, where E_Λ denotes the expectation operator with respect to $\Lambda.\mu$ and $D_t \Lambda$ is the Lebesgue density of the absolutely continuous map $t \to \nabla \Lambda(t,x)$. From the relation (6.12), it follows that α is a function of T, hence we have obtained the strong solution of the Monge-Ampère equation. Let us announce all this as

Theorem 6.1. *Assume that K and L are upper and lower bounded densities, let T be the transport map constructed in Theorem 4.1. Then T is also the strong solution of the Monge-Ampère equation in the Itô*

sense, namely

$$E[\Lambda^2]\, L \circ T \exp\left\{-\int_0^1 \dot{\alpha}_s dT_s - \frac{1}{2}\int_0^1 |\dot{\alpha}_s|^2 ds\right\} = K\,,$$

μ-almost surely, where α is given with (6.12).

References

[1] P. Appell: "Mémoire sur déblais et les remblais des systèmes continus ou discontinus". Mémoires présentées par divers savants à l'Académie des Sciences de l'Institut de France. Paris, I. N. **29**, 1-208, 1887.

[2] P. Appell: "Le problème géométrique des déblais et des remblais". Mémorial des Sciences Mathématiques, fasc. **XXVII**, Paris, 1928.

[3] P. J. Bickel and D. A. Freedman: "Some asymptotic theory for the bootstrap". The Annals of Statistics, Vol. **9**, No. 6, 1196-1217, 1981.

[4] Y. Brenier: "Polar factorization and monotone rearrangement of vector valued functions". Comm. pure Appl. Math, **44**, 375-417, 1991.

[5] L. A. Caffarelli: "The regularity of mappings with a convex potential". Jour. Amer. Math. Soc., **5**, 99-104, 1992.

[6] C. Dellacherie and P. A. Meyer: *Probabilités et Potentiel, Ch. I à IV.* Paris, Hermann, 1975.

[7] N. Dunford and J. T. Schwartz: *Linear Operators 2*, Interscience 1963.

[8] X. Fernique: "Extension du théorème de Cameron-Martin aux translations aléatoires", Comptes Rendus Mathématiques, Vol. **335**, Issue 1, 65-68, 2002.

[9] X. Fernique: "Comparaison aux mesures gaussiennes, espaces autoreproduisants. Une application des propriétés isopérimétriques". Preprint.

[10] D. Feyel and A. de La Pradelle: "Capacités gaussiennes". Annales de l'Institut Fourier, t.41, f.1, 49-76, 1991.

[11] D. Feyel and A. S. Üstünel: "The notion of convexity and concavity on Wiener space". Journal of Functional Analysis, **176**, 400-428, 2000.

[12] D. Feyel and A. S. Üstünel: "Transport of measures on Wiener space and the Girsanov theorem". Comptes Rendus Mathématiques, Vol. **334**, Issue 1, 1025-1028, 2002.

[13] D. Feyel and A. S. Üstünel: "Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space". Preprint, 2003.

[14] W. Gangbo and R. J. McCann: "The geometry of optimal transportation". Acta Mathematica, **177**, 113-161, 1996.

[15] K. Itô and M. Nisio: "On the convergence of sums of independent Banach space valued random variables". Osaka Journ. Math. **5**, 35-48, 1968.

[16] L. V. Kantorovitch: "On the transfer of masses". Dokl. Acad. Nauk. SSSR **37**, 227-229, 1942.

[17] K. Marton: "Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration". Annals of Probability, **24**, no.2, 857-866, 1996.

[18] R. J. McCann: "Existence and uniqueness of monotone measure-preserving maps". Duke Math. Jour., **80**, 309-323, 1995.

[19] R. J. McCann: "A convexity principle for interacting gases". Advances in Mathematics, **128**, 153-179, 1997.

[20] G. Monge: "Mémoire sur la théorie des déblais et des remblais". Histoire de l'Académie Royale des Sciences, Paris, 1781.

[21] F. Otto and C. Villani: "Generalization of an inequality by Talagrand and links with logarithmic Sobolev inequality". Journal of Functional Analysis, **173**, 361-400, 2000.

[22] T. Rockafellar: *Convex Analysis*. Princeton University Press, Princeton, 1972.

[23] V. N. Sudakov: "Geometric problems in the theory of infinite dimensional probability distributions". Proc. Steklov Inst. Math., **141**, 1-178, 1979.

[24] M. Talagrand: "Transportation cost for Gaussian and other product measures". Geom. Funct. Anal., **6**, 587-600, 1996.

[25] H. Tanaka "An inequality for a functional of probability distributions and its applications to Kac's one-dimensional model of a Maxwellian gas". Z. für Wahrsch. verw. Geb., **27**, 47-52, 1973.

[26] H. Tanaka "Certain functional of probability measures on Hilbert spaces". Hiroshima Math. Jour., **6**, 421-428, 1976.

[27] A. S. Üstünel: "Representation of distributions on Wiener space and Stochastic Calculus of Variations". Journal of Functional Analysis, **70**, 126-139, 1987.

[28] A. S. Üstünel: *Introduction to Analysis on Wiener Space*. Lecture Notes in Math. Vol. **1610**. Springer, 1995.

[29] A. S. Üstünel and M. Zakai: *Transformation of Measure on Wiener Space*. Springer Monographs in Mathematics. Springer Verlag, 1999.

Denis Feyel
Université d'Evry-Val-d'Essone,
91025 Evry Cedex,
France.
E-mail: `feyel@maths.univ-evry.fr`

Ali Süleyman Üstünel
ENST,
Dépt. Infres,
46, rue Barrault, 75634 Paris Cedex 13,
France.
E-mail: `ustunel@enst.fr`