Monge-Ampère Equation on Wiener Space

D. Feyel and A. S. Üstünel

Abstract

In this work we study the Monge-Ampère equation in the frame of infinite dimensional Fréchet spaces equipped with a Gauss measure, polar factorization of the mappings which transform a spread measure to another one in terms of the measure transportation of Monge-Kantorovich.

Contents

1 Introduction

In 1781, G. Monge has launched his famous problem, which can be expressed in terms of the modern mathematics as follows: given two probability measures ρ and ν on $\mathbb{R}^n$, find the map $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $T\rho = \nu$ [1] and T is also the

[1] $T\rho$ means the image of the measure ρ under the map T

solution of the minimization problem

$$\inf_U \left\{ \int_{\mathbb{R}^n} c(x, U(x))\rho(dx) \right\} , \tag{1.1}$$

where the infimum is taken between all the maps $U : \mathbb{R}^n \to \mathbb{R}^n$ such that $U\rho = \nu$ and where $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ is a positive, measurable function, called usually the cost function. In the original problem of Monge, the cost function $c(x, y)$ was $|x - y|$, later other costs have been considered. Between them, the most popular one which is also abundantly studied, is the case where $c(x, y) = |x - y|^2$.After several tentatives (cf., [1, 2]), in the 1940's this highly nonlinear problem of Monge has been reduced to a linear problem by Kantorovitch, cf.[12], in the following way: let $\Sigma(\rho, \nu)$ be the set of probability measures on $\mathbb{R}^n \times \mathbb{R}^n$, whose first marginals are ρ and the second marginals are ν. Find the element(s) of $\Sigma(\rho, \nu)$ which are the solutions of the minimization problem:

$$\inf_{\beta \in \Sigma(\rho,\nu)} \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) d\beta(x, y) \right\} . \tag{1.2}$$

It is obviuous that $\Sigma(\rho, \nu)$ is a convex, compact set under the weak*-topology of measures, hence, in case, the cost function c has some regularity properties, like being lower semi-continuous, this problem would have solutions. If any one of them is supported by the graph of a map $T : \mathbb{R}^n \to \mathbb{R}^n$, then obviously, T will be also a solution of the original problem of Monge 1.1. Since that time, the problem (1.2) is called the Monge-Kantorovitch problem (MKP). The program of Kantorovitch has been followed by several people and a major contribution has been done by Sudakov [18]. In the early 90's there has been another impetus to this problem, cf., [4], where it has been proven the importance of the convex functions in the solutions of the MKP and of the problem of Monge (cf., also [13, 14]).

In [9], we have solved the MKP and the problem of Monge in the infinite dimensional case, where the space on which the measures are defined is a Polish space into which is injected densely and continuously a Hilbert space H, which is called as the Cameron-Martin space in reference to the Gaussian case. The cost function that we have chosen was $c(x, y) = |x - y|_H^2$, where $| \cdot |_H$ denotes the Euclidean norm of H. Because of this choice, in comparison to the finite dimensional space, the situation becomes very singular, since, in general, the space H is of zero measure with respect to almost all the reasonable measures for which one can expect to have solutions for the aforementioned problems of Monge and MKP. On the other hand, due to the potential applications to several problems of stochastic analysis and physics, this cost function is particularly important. We have also given several applications of these results, for example we have proven a factorization result of the absolutely continuous transformations of the Wiener measure, the solutions of the Monge-Ampère equation, etc.

This paper is devoted to the further developement of the subject, in particular, we give a much more general result of the factorization of the vector fields which map a probability measure on W to another one such that one of them is spread and (cf. the preliminaries) the two measures are at finite Wasserstein distance from each other. As an example we treat in detail the case of the Gaussian measures. One of the main contributions in this work is the proof of the existence of the functional analytic (or strong) solutions of the Monge-Ampère equation for the H-log-concave densities. In [9], we have studied the Monge-Ampère equation for the upper and lower bounded densities with respect to the Wiener measure. The main difficulty in this case stems from the lack of regularity of the transport potentials, in fact we know only that these functions are in the Sobolev space $\mathbb{D}_{2,1}$, i.e., they have only first order Sobolev derivatives. However, to write the Gaussian Jacobian, we need them to be second order Sobolev differentiable. This difficulty, which exists also in the finite dimensional case is solved in this latter case with the help of the Alexandroff differentiability of the convex functions. In the infinite dimensional case the situation is worse, in fact the transport potentials are not in general convex, nor H-convex (which is a more reasonable requirement, cf. [7]), but only 1-convex in the Cameron-Martin space direction as explained in the next sections. As it is shown in Section 5, the only exception to this lack of regularity arises in the case where the target measure has an H-log-concave density. In this case, using the finite dimensional results of Caffarelli [6], we show that the transport potential has a second order derivative as an operator valued map and then using some celebrated identity of Wiener space analysis, we also prove that this second derivative takes its values in the space of Hilbert-Schmidt operators, hence we can write the corresponding Jacobian which includes the modified Carleman-Fredholm determinant, cf. [22] and finally we prove that the transport potential is the unique 1-convex solution of the Monge-Ampère equation. In Section 6 we show that all these difficulties disappear if we use the natural Ito Calculus and we can calculate the Jacobian (cf. Theorem 6.1) using the natural Brownian motion which is associated to the solution of the Monge problem. In fact, with Itô parametrization, the complications are absorbed by the filtrations of forward and backward transport processes (i.e., maps).

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 V_n be the σ-algebra generated by π_{F_n}, then for any $f \in L^p(\mu)$, the martingale sequence $(E[f|V_n], n \geq 1)$ converges to f (strongly if $p < \infty$) in $L^p(\mu)$. Observe that the function $f_n = E[f|V_n]$ 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 [21]. 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.[21]).

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.[21]). 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). Due to the Meyer inequalities (cf., for instance [21]), the norms defined by

$$\|\varphi\|_{p,k} = \|(I + \mathcal{L})^{k/2}\varphi\|_{L^p(\mu)} \tag{2.3}$$

are equivalent to the norms defined by the iterates of the Sobolev derivative ∇. This observation permits us to identify the duals of the space $\mathbb{D}_{p,k}(\Phi); p > 1$, $k \in \mathbb{N}$ by $\mathbb{D}_{q,-k}(\Phi')$, with $q^{-1} = 1 - p^{-1}$, where the latter space is defined by replacing k in (2.3) 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 [21]). The following assertion is useful: assume that $(Z_n, n \geq 1) \subset \mathbb{D}'$ converges to Z in $\mathbb{D}'$, assume further that each each Z_n is a probability measure on W, then Z is also a probability and $(Z_n, n \geq 1)$ converges to Z in the weak topology of measures. In particular, a lower bounded distribution (in the sense that there exists a constant $c \in \mathbb{R}$ such that $Z + c$ is a positive distribution) is a (Radon) measure on W, c.f. [21].

A measurable function $f : W \to \mathbb{R} \cup \{\infty\}$ is called H-convex (cf.[7]) if

$$h \to f(x + h)$$

is convex μ-almost surely, i.e., if for any $h, k \in H$, $s, t \in [0, 1]$, $s + t = 1$, we have

$$f(x + sh + tk) \leq sf(x + h) + tf(x + k)\,,$$

almost surely, where the negligeable set on which this inequality fails may depend on the choice of s, h and of k. We can rephrase this property by saying that $h \to f(x + h)$ is an $L^0(\mu)$-valued convex function on H. f 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 all these notions are compatible with the μ-equivalence classes of random variables thanks to the Cameron-Martin theorem. It is proven in [7] 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. We define similarly the notion of H-concave and H-log-concave functions. In particular, one can prove that, for any H-log-concave function f on W, $P_t f$ and $E[f|V_n]$ are again H-log-concave [7].

3 Monge-Kantorovitch problem

Let us recall the definition of the Monge-Kantorovitch problem in our case:

Definition 3.1 *Let ρ and ν be two probability measures on W, let also $\Sigma(\rho, \nu)$ be the convex subset of the probability measures on the product space $W \times W$ whose first marginal is ρ and the second one is ν. The Monge-Kantorovitch problem for the couple (ρ, ν) consists of finding a measure $\gamma \in \Sigma(\rho, \nu)$ which realizes the following infimum:*

$$d_H^2(\rho, \nu) = \inf_{\beta \in \Sigma(\rho,\nu)} \int_{W \times W} |x - y|_H^2 d\beta(x, y)\,.$$

The function $c(x, y) = |x - y|_H^2$ is called the cost function.

Remark 3.1 *Note that the cost function is not continuous with respect to the product topology of $W \times W$ and it takes the value ∞ very often for the most notable measures, e.g., when ρ and ν are absolutely continuous with respect to the Wiener measure μ.*

The proof of the next theorem, for which we refer the reader to [9], 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 measurable section-selection theorem.

Theorem 3.1 (General case) *Suppose that ρ and ν are two probability measures on W such that*

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

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

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

and for $\pi_n^\perp\rho$-almost almost all $x_n^\perp$, the map $u \to u + \xi(u + x_n^\perp)$ 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

$$\begin{aligned} 1 &= \gamma\left\{(x, y) \in W \times W : T \circ S(y) = y\right\} \\ &= \gamma\left\{(x, y) \in W \times W : S \circ T(x) = x\right\} \,. \end{aligned}$$

In particular we have

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

Remark 3.2 *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 $n-1$.*

Any probability measure satisfying the hypothesis of Theorem 3.1 is called a spread measure. Namely,

Definition 3.2 *A probability measure m on $(W,\mathcal{B}(W))$ is called a spread measure if there exists a sequence of finite dimensional regular projections $(\pi_n, n\geq 1)$ converging to I_H such that the regular conditional probabilities $m(\,\cdot\,|\pi_n^\perp = x_n^\perp)$ concentrated in the n-dimensional spaces $\pi_n(W)+x_n^\perp$ vanishe on the sets of Hausdorff dimension $n-1$ for $\pi_n^\perp(m)$-almost all $x_n^\perp$ and for any $n\geq 1$.*

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 3.1:

Theorem 3.2 (Gaussian case) *Let ν be the measure $d\nu = L d\mu$, 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 $\varphi\in \mathbb{D}_{2,1}$ and a partially 1-convex function $\psi\in L^2(\nu)$, both are unique upto a constant, called Monge-Kantorovitch potentials, such that*

$$\varphi(x)+\psi(y)+\frac{1}{2}|x-y|_H^2 \geq 0$$

for all $(x,y)\in W\times W$ and that

$$\varphi(x)+\psi(y)+\frac{1}{2}|x-y|_H^2 = 0$$

γ-almost everywhere. The map $T = I_W + \nabla\varphi$ 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 $T^{-1} = I_W + \eta$ such that $T^{-1}\nu = \mu$, $\eta \in L^2(\nu)$ 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 3.3 *By the partial 1-convexity we mean that $y_F \to \psi(y_F + y_F^\perp)$ is $\nu(\cdot|\pi_F^\perp = y_F^\perp)$-almost surely 1-convex on any regular, finite dimensional subspace F, where π_F is the (regular) projection corresponding to F, $y_F = \pi_F(y)$ and $y_F^\perp = y - y_F$. 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 3.4 *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 $\varphi_n \in \mathbb{D}_{2,1}$ is the function constructed in Theorem 3.2, corresponding to L_n, then, using the inequality (cf., [9])*

$$d_H^2(\mu, \nu) \leq 2E[L \log L] ,$$

we can prove that the sequence $(\varphi_n, n \in \mathbb{N})$ converges to φ in $\mathbb{D}_{2,1}$.

4 Polar factorization of mappings between spread measures

In [9] we have proved the polar factorization of the mappings $U : W \to W$ such that the Wasserstein distance between $U\mu$ and the Wiener measure μ, denoted by $d_H(\mu, U\mu)$, is finite. We have also studied the particular case where U is a perturbations of identity, i.e., it is the form $I_W + u$, where u maps W to the Cameron-Martin space H. In this section we shall generalize this results in the frame of spread measures.

Theorem 4.1 *Assume that ρ and ν are spread measures with $d_H(\rho, \nu) < \infty$ and that $U\rho = \nu$, for some measurable map $U : W \to W$. Let T be the optimal transport map sending ρ to ν, whose existence and uniqueness is proven in Theorem 3.1. Then $R = T^{-1} \circ U$ is a ρ-rotation (i.e., $R\rho = \rho$) and $U = T \circ R$, morover, if U is a perturbation of identity, then R is also a perturbation of identity. In both cases, R is the ρ-almost everywhere unique minimal ρ-rotation in the sense that*

$$\int_W |U(x) - R(x)|_H^2 d\rho(x) = \inf_{R' \in \mathcal{R}} \int_W |U(x) - R'(x)|_H^2 d\rho(x) , \qquad (4.4)$$

where $\mathcal{R}$ denotes the set of ρ-rotations.

Proof: Let T be the optimal transportation of ρ to ν whose existence and uniqueness follows from Theorem 3.1. The unique solution γ of the Monge-Kantorovitch problem for $\Sigma(\rho,\nu)$ can be written as $\gamma=(I\times T)\rho$. Since ν is spread, T is invertible on the support of ν and we have also $\gamma=(T^{-1}\times I)\nu$. In particular $R\rho=T^{-1}\circ U\rho=T^{-1}\nu=\rho$, hence R is a rotation. Let R' be another rotation in $\mathcal{R}$, define $\gamma'=(R'\times U)\rho$, then $\gamma'\in\Sigma(\rho,\nu)$ and the optimality of γ implies that $J(\gamma)\le J(\gamma')$, besides we have

$$\begin{aligned}\int_W |U(x)-R(x)|_H^2 d\rho(x) &= \int_W |U(x)-T^{-1}\circ U(x)|_H^2 d\rho(x)\\ &= \int_W |x-T^{-1}(x)|_H^2 d\nu(x)\\ &= \int_W |T(x)-x|_H^2 d\rho(x)\\ &= J(\gamma)\,.\end{aligned}$$

On the other hand

$$J(\gamma')=\int_W |U(x)-R'(x)|_H^2 d\rho(x)\,,$$

hence the relation (4.4) follows. Assume now that for the second rotation $R'\in\mathcal{R}$ we have the equality

$$\int_W |U(x)-R(x)|_H^2 d\rho(x)=\int_W |U(x)-R'(x)|_H^2 d\rho(x)\,.$$

Then we have $J(\gamma)=J(\gamma')$, where γ' is defined above. By the uniqueness of the solution of Monge-Kantorovitch problem due to Theorem 3.1, we should have $\gamma=\gamma'$. Hence $(R\times U)\rho=(R'\times U)\rho=\gamma$, consequently, we have

$$\int_W f(R(x),U(x))d\rho(x)=\int_W f(R'(x),U(x))d\rho(x)\,,$$

for any bounded, measurable map f on $W\times W$. This implies in particular

$$\int_W (a\circ T\circ R)\ (b\circ U)d\rho=\int_W (a\circ T\circ R')\ (b\circ U)d\rho$$

for any bounded measurable functions a and b. Let $U'=T\circ R'$, then the above expression reads as

$$\int_W a\circ U\ b\circ U d\rho=\int_W a\circ U'\ b\circ U d\rho\,.$$

Taking $a=b$, we obtain

$$\int_W (a\circ U)\,(a\circ U')\,d\rho=\|a\circ U\|_{L^2(\rho)}\|a\circ U'\|_{L^2(\rho)}\,,$$

for any bounded, measurable a. This implies that $a \circ U = a \circ U'$ ρ-almost surely for any a, hence $U = U'$ i.e, $T \circ R = T \circ R' \rho$-almost surely. Let us denote by S the left inverse of T whose existence follows from Theorem 3.1 and let $D = \{x \in W : S \circ T(x) = x\}$. Since $\rho(D) = 1$ and since R and R' are ρ-rotations, we have also

$$\rho\left(D \cap R^{-1}(D) \cap R'^{-1}(D)\right) = 1 .$$

Let $x \in W$ be any element of $D \cap R^{-1}(D) \cap R'^{-1}(D)$, then

$$\begin{aligned} R(x) &= S \circ T \circ R(x) \\ &= S \circ T \circ R'(x) \\ &= R'(x) , \end{aligned}$$

consequently $R = R'$ on a set of full ρ-measure.. $\square$

Let us give another result of interest as an application of these factorization results: it is important to have as much as information about the measures and the tranformations which induce them in the setting of Girsanov Theorem, cf. [22] and the references there. The problem which we propose is the following: assume that, in the case of the Wiener measure, we have a density L with $d_H(\mu, L{\cdot}\mu) < \infty$, hence from Theorem 3.1, a map $T : W \to W$, which is the optimal transport map corresponding to the solution of MKP in $\Sigma(\mu, L \cdot \mu)$. Since the target measure is also spread, the map T possesses a left inverse S such that $S \circ T = I_W$ μ-almost surely. Assume now that the transformation T has a Girsanov density, i.e., $\lambda \in L^1_+(\mu)$, with $E[\lambda] = 1$ and that

$$\int f \circ T \lambda \, d\mu = \int f \, d\mu ,$$

for any $f \in C_b(W)$. We can now prove:

Theorem 4.2 *Let T be as explained above, assume moreover that*

$$d_H(\lambda \cdot \mu, \mu) < \infty ,$$

then T has also a right inverse, i.e., T is invertible μ-almost everywhere.

Proof: Denote by $\Theta : W \to W$ the optimal transportation map corresponding to the solution of MKP in $\Sigma(\mu, \lambda{\cdot}\mu)$. Note that both of the measures $(T \times I_W)(\lambda{\cdot}\mu)$ and $(I_W \times \Theta)\mu$ belong to $\Sigma(\mu, \lambda \cdot \mu)$. By the uniqueness of the solutions of MKP, they are equal, hence, for any $a, b \in C_b(W)$, we have

$$\int a(T(x)) \, b(x)\lambda(x) d\mu(x) = \int a(x) b(\Theta(x)) d\mu(x) . \tag{4.5}$$

Since $\Theta(\mu) = \lambda \cdot \mu$, the equality (4.5) can also be written as

$$\int a(T \circ \Theta(x))\, b(\Theta(x)) d\mu(x) = \int a(x) b(\Theta(x)) d\mu(x)\,. \tag{4.6}$$

Since, as T, the map Θ has also a left inverse, the sigma algebra generated by Θ is equal to the Borel sigma algebra of W, consequently, the relation (4.6) implies that

$$a \circ T \circ \Theta = a\,,$$

μ-almost surely, for any $a \in C_b(W)$. Therefore we have

$$\mu\left(\{x \in W : T \circ \Theta(x) = x\}\right) = 1\,,$$

since T has already a left inverse, the proof is completed. □

4.1 Application to Gaussian measures

Let us give an example of the above results: Assume that $\rho = \mu$, i.e., the Wiener measure and let K be a Hilbert-Schmidt operator on H. Assume that the Carleman-Fredholm determinant $\det_2(I_H + K)$ is different than zero, hence the operator $I_H + K : H \to H$ is invertible. Moreover, it follows from the general theory that $I_H + K$ has a unique polar decomposition as $I_H + K = (I_H + \bar{K})(I_H + A)$, where $I_H + A$ is an isometry[2] and $I_H + \bar{K}$ is a symmetric, positive operator. Note that $\bar{K}$ is compulsorily Hilbert-Schmidt. Let us now define $U : W \to W$ as $U(x) = x + \delta K(x)$, where $\delta K(x)$ is the H-valued divergence, defined by $(\delta K(x), h)_H = \delta(K^* h)(x)$. Then it is known that the measure $U\mu$ is absolutely continuous with respect to μ, in fact $U\mu$ is even equivalent to μ since $|\Lambda_K| \neq 0$ μ-almost surely, where

$$\Lambda_K = \det_2(I_H + K) \exp\left\{\delta^2(K) - \frac{1}{2}|\delta K|_H^2\right\}\,.$$

Besides we have

$$L = \frac{dU\mu}{d\mu} = \frac{1}{|\Lambda_K| \circ V}\,,$$

where V is the inverse of U, whose existence follows from the invertibility of $h \to h + \delta(K)(x) + Kh$ on H, cf. [22]. Consequently,

$$E[L \log L] = -E[\log |\Lambda_K|] < \infty\,,$$

hence $d_H(\mu, U\mu) < \infty$. We shall prove that the polar factorization of U is given by

$$U = (I_W + \delta \bar{K}) \circ (I_W + \delta A)\,.$$

[2] A satisfies the relation $A + A^* + A^* A = 0$.

In fact, it follows from Theorem B.6.4 of [22], that

$$\begin{aligned}(I_W+\delta\bar K)\circ(I_W+\delta A) &= I_W+\delta\bar K+\delta A+\delta(\bar K A)\\ &= I_W+\delta(\bar K+A+\bar K A)\\ &= I_W+\delta K\,.\end{aligned}$$

Besides $\nabla^2\delta^2\bar K=2\bar K$, and since $I_H+\bar K$ is a positive operator, the Wiener map $\frac{1}{2}\delta^2\bar K$ is 1-convex, consequently, $T=I_W+\delta\bar K$ is the transport map and $I_W+\delta A$ is the unique rotation whose existence is proven in Theorem 4.1. The Kantorovitch potentials φ and ψ of Theorem 3.2 can be chosen as

$$\varphi(x)=\frac{1}{2}\delta^2\bar K(x)$$

for T and

$$\psi(x)=-\frac{1}{2}\delta((I_H+\bar K)^{-1}\bar K)(x)$$

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

Remark 4.1 Let us denote by $P_{\ker\delta}$ the projection operator from $\mathbb{D}'(H)$ to the kernel of the divergence operator δ. Then, we have the following identity:

$$P_{\ker\delta}\left(\delta((I_H+\bar K)A)\right)=\delta\hat K-\delta\bar K\,,$$

where $\hat K$ denotes the symmetrization of K. This shows that the polar decomposition and the Helmholtz decomposition are different in general.

We can also calculate the Monge-Kantorovitch potential function for the singular case as follows: assume that ν is a zero mean Gaussian measure on W such that $d_H(\mu,\nu)<\infty$. Then there exists a bilinear form q on $W^\star$ such that

$$\int_W e^{i\langle\alpha,x\rangle}d\nu(x)=\exp-\frac{1}{2}q(\alpha,\alpha)\,,$$

for any $\alpha\in W^\star$. On the other hand, from Theorem 3.2, there exists a $\varphi\in\mathbb{D}_{2,1}$, which is 1-convex, such that $T\mu=(I_W+\nabla\varphi)\mu=\nu$. Hence, rewriting the above relation with T, we obtain:

$$\int_W e^{i\langle t\alpha,T(x)\rangle}d\mu(x)=\exp-\frac{t^2}{2}q(\alpha,\alpha)\,,\tag{4.7}$$

for any $t\in\mathbb{R}$ and $\alpha\in W^\star$. Taking the derivative of both sides twice at $t=0$, we obtain

$$\begin{aligned}q(\alpha,\alpha) &= |\tilde\alpha|_H^2+E\left[(\nabla\varphi,\tilde\alpha)_H^2\right]+2E\left[(\nabla\varphi,\tilde\alpha)_H\delta\tilde\alpha\right]\\ &= |\tilde\alpha|_H^2+E\left[(\nabla\varphi\otimes\nabla\varphi,\tilde\alpha\otimes\tilde\alpha)_2\right]+2E\left[(\nabla^2\varphi,\tilde\alpha\otimes\tilde\alpha)_2\right]\,,\end{aligned}$$

where $\tilde{\alpha}$ denotes the image of α under the injection $W^\star \hookrightarrow H$.Note that, here, $\nabla^2\varphi$ is to be interpreted as a distribution. Denote by M the Hilbert-Schmidt operator defined by

$$M = E\left[\nabla\varphi\otimes\nabla\varphi\right] + 2E\left[\nabla^2\varphi\right] .$$

We have

$$q(\alpha,\alpha) = ((I_H + M)\tilde{\alpha},\tilde{\alpha})_H .$$

Let $I_H + N$ be the positive square root of the (positive) operator $I_H + M$, then N is a symmetric Hilbert-Schmidt operator. Define

$$\varphi = \frac{1}{2}\delta^2 N .$$

Evidently φ is a 1-convex element of $\mathbb{D}_{2,1}$, moreover the map T defined by $T = I_W + \nabla\varphi = I_W + \delta N$ satisfies the identity (4.7), hence T is the unique solution of the Monge problem and $(I_W \times T)\mu$ is the unique solution of MKP for $\Sigma(\mu,\nu)$.

5 Strong solutions of the Monge-Ampère equation for H-log-concave densities

Assume that $L \in \mathbb{L}^1_{+,1}(\mu)$ is of the form

$$L = \frac{1}{E\left[e^{-f}\right]}e^{-f} ,$$

where f is an H-convex function in some $L^p(\mu)$, $p > 1$. We assume that $f \geq -\alpha$ almost surely, for some $\alpha \in \mathbb{R}_+$. Denote by $\varphi \in \mathbb{D}_{2,1}$ the potential of the transport problem between μ and $L\cdot\mu$ which is a 1-convex function. This means that the mapping defined by $T = I_W + \nabla\varphi$ satisfies $T\mu = L\cdot\mu$ and $(I_W\times T)\mu$ is the unique solution of the Monge-Kantorovitch problem with the singular quadratic cost function $c(x,y) = |x-y|^2_H$. Let $\Lambda = 1/L\circ T$, we know that $T^{-1}\mu = \Lambda\cdot\mu$ where $T^{-1} = I_W + \nabla\psi$ such that $\psi \in \mathbb{D}_{2,1}$ is also defined uniquely. Let $L_n = E[L|V_n]$, where V_n is the sigma algebra generated by the first n elements of an orthonormal basis $(e_n, n \geq 1)$ of H. It follows from [7], that L_n is of the form $\frac{1}{c}e^{-f_n}$, where f_n is an H-convex function on W and $c = E[e^{-f}]$. We denote by φ_n, Λ_n, ψ_n the maps associated to L_n, i.e., $T_n = I_W + \nabla\varphi_n$ maps μ to the measure $L_n\cdot\mu$ and $S_n = I_W + \nabla\psi_n$ maps $L_n\cdot\mu$ to μ. Besides, from [6], $\nabla\varphi_n$ is a 1-Lipschitz map, i.e.,

$$|\nabla\varphi_n(x) - \nabla\varphi_n(y)| \leq |x-y| ,$$

for any $x, y \in \mathbb{R}^n$, here it is remarkable that the Lipschitz constant is one and it is independent of the dimension of the underlying space. Hence $\mathcal{L}\varphi_n$ is a well-defined element of $L^2(\mu)$, $|\nabla\varphi|^2_H$ is exponentially integrable, i.e., there exists some

$t > 0$ such that

$$\sup_n E\left[\exp t|\nabla\varphi_n|_H^2\right] < \infty\,. \tag{5.8}$$

It follows in particular that $(\varphi_n, n \geq 1) \subset \mathbb{D}_{p,2}$ and it converges to φ in $\mathbb{D}_{p,1}$ for any $p \geq 1$, cf., [9]. Moreover, from a result of McCann [14], we have

$$\Lambda_n = \det{}_2(I_H + \nabla^2\varphi_n)\exp\left\{-\mathcal{L}\varphi_n - \frac{1}{2}|\nabla\varphi_n|_H^2\right\}.$$

Since $\Lambda_n = 1/L_n \circ T_n$, the sequence $(\Lambda_n, n \geq 1)$ is lower bounded. Hence $(-\log\Lambda_n, n \geq 1)$ is upper bounded, besides

$$\begin{aligned} E[|\log\Lambda_n|^p] &\leq C_pE[|f_n \circ T_n|^p] + D_p \\ &= C_pE[|f_n|^pL_n] + D_p \\ &\leq C_pe^{p\alpha}E[|f|^p] + D_p\,, \end{aligned}$$

where C_p and D_p are some constants. Since $(-\log\Lambda_n, n \geq 1)$ converges in $\mathbb{L}^0(\mu)$ to $-\log\Lambda$, it follows from the dominated convergence theorem that $(\log\Lambda_n, n \geq 1)$ converges to $\log\Lambda$ in $\mathbb{L}^p(\mu)$. Therefore

$$-\log\det{}_2(I_H + \nabla^2\varphi_n) + \mathcal{L}\varphi_n + \frac{1}{2}|\nabla\varphi_n|_H^2 \to -\log\Lambda$$

in $\mathbb{L}^p(\mu)$. Since $(\varphi_n, n \geq 1)$ converges to φ in $\cap_p\mathbb{D}_{p,1}$, the sequence $(Z_n, n \geq 1)$, defined by

$$Z_n = -\log\det{}_2(I_H + \nabla^2\varphi_n) + \mathcal{L}\varphi_n\,,$$

converges in $\mathbb{L}^p(\mu)$ to some $Z \in \mathbb{L}^p(\mu)$. Again by the convergence of $(\varphi_n, n \geq 1)$, the sequence $(\mathcal{L}\varphi_n, n \geq 1)$ converges to the measure $\mathcal{L}\varphi$ in $\mathbb{D}_{2,-1}$ (cf. [9]), consequently the sequence $(\log\det_2(I_H + \nabla^2\varphi_n), n \geq 1)$ converges to some $D = D(\varphi)$ in $\mathbb{D}'$. Since $Z = \mathcal{L}\varphi - D(\varphi)$ and $\mathcal{L}\varphi$ are measures, $D(\varphi)$ should be a measure, besides Z is absolutely continuous with respect to μ (it is a random variable), hence $\mathcal{L}_s\varphi - D_s(\varphi) = 0$, where the subscript "s" denotes the singular part of the measure $D(\varphi)$. Consequently we have $Z = \mathcal{L}_a\varphi - D_a(\varphi)$, where the subscript "$a$" denotes the absolutely continuous part of the corresponding measure. Therefore we have

$$\begin{aligned} \Lambda &= \lim \Lambda_n \\ &= \exp\left\{D_a(\varphi) - \mathcal{L}_a\varphi - \frac{1}{2}\,|\nabla\varphi|_H^2\right\}. \end{aligned}$$

In fact we have a much better result of regularity:

Theorem 5.1 *Assume further that $f \in \mathbb{D}_{2,1}$, then $\varphi \in \mathbb{D}_{2,2}$, in particular*

$$\mathcal{L}_a\varphi = \mathcal{L}\varphi \in L^2(\mu)$$

and $\det_2(I_H + \nabla^2\varphi)$ is a well-defined function.

Proof: L_n is μ-a.s. strictly positive by the hypothesis that we have done for L. Consequently, the operator $I+\nabla\varphi_n(x)$ is almost surely invertible. Besides, using the commutation relation between the Gaussian divergence and an absolutely continuous transformation of the Wiener measure (cf. [22], Appendix B) and the relation $\delta\circ\nabla=\mathcal{L}$, we get

$$\mathcal{L}\psi_n\circ T_n=\delta(\nabla\psi_n\circ T_n)+(\nabla\psi_n\circ T_n,\nabla\varphi_n)_H+\text{ trace }(\nabla^2\psi_n\circ T_n\cdot\nabla^2\varphi_n)\,. \quad (5.9)$$

It is easy to see that

$$\text{trace }(\nabla^2\psi_n\circ T_n\cdot\nabla^2\varphi_n)=-\text{ trace }\left((I+\nabla^2\varphi_n)^{-1}\cdot(\nabla^2\varphi_n)^2\right)\,.$$

Taking the expectation of both sides of (5.9) with respect to μ, we have

$$E\left[\text{ trace }\left((I+\nabla^2\varphi_n)^{-1}\cdot(\nabla^2\varphi_n)^2\right)\right]=E[|\nabla\varphi_n|_H^2]-E[\mathcal{L}\psi_n\,L_n]\,.$$

Since $(L_n,n\geq 1)$ is uniformly essentially bounded by some $K>0$, we have

$$\begin{aligned}E[\mathcal{L}\psi_n\,L_n] &= E[(\nabla\psi_n,\nabla L_n)_H]\\ &= -E[(\nabla\psi_n,\nabla f_n)_H L_n]\\ &\leq K\,\|\nabla\psi_n\|_{L^2(\mu,H)}\|f\|_{2,1}\,.\end{aligned}$$

Moreover, from the Young inequality

$$E[|\nabla\psi_n|_H^2]=E[|\nabla\varphi_n|_H^2\,\Lambda_n]\leq E[\varepsilon^{-1}\Lambda_n\log\Lambda_n]+E\left[\exp\varepsilon|\nabla\varphi_n|_H^2\right]\,,$$

which is uniformly bounded with respect to n since $\|\nabla^2\varphi_n\|_{\text{op}}\leq 1$. Consequently

$$\sup_n E\left[\text{ trace }\left((I+\nabla^2\varphi_n)^{-1}\cdot(\nabla^2\varphi_n)^2\right)\right]<\infty\,.$$

Recalling that $\|I_H+\nabla^2\varphi_n\|_{\text{op}}\leq 1$ almost surely, we finally get

$$\begin{aligned}\sup_n E\left[\text{ trace }(\nabla^2\varphi_n)^2\right] &= \sup_n E[\|\nabla^2\varphi_n\|_2^2]\\ &\leq \sup_n E\left[\left\|(I_H+\nabla^2\varphi_n)^{-1/2}\nabla^2\varphi_n\right\|_2^2\right]\\ &= \sup_n E\left[\text{ trace }\left((I+\nabla^2\varphi_n)^{-1}\cdot(\nabla^2\varphi_n)^2\right)\right]<\infty\,.\end{aligned}$$

This implies that $(\nabla^2\varphi_n,n\geq 1)$ is bounded in the space Hilbert-Schmidt valued Wiener maps $L^2(\mu,H\otimes H)$, since $(\varphi_n,n\geq 1)$ converges to φ in $\mathbb{D}_{2,1}$, φ should be in $\mathbb{D}_{2,2}$ and the other claims are now immediate. $\square$

Corollary 5.1 *Let λ be the function defined as*

$$\lambda = \det{}_2(I_H + \nabla^2\varphi)\exp\left\{-\mathcal{L}\varphi - \frac{1}{2}|\nabla\varphi|_H^2\right\}.$$

Then λ is a sub-solution of the Monge-Ampère equation in the sense that

$$E[g \circ T\,\lambda] \leq E[g]\,, \tag{5.10}$$

for any positive, measurable function g. In particular

$$\lambda \leq \Lambda$$

almost surely.

Proof: Let $(e_n, n \geq 1) \subset W^*$ be a complete, orthonormal basis of H, denote by V_n the sigma algebra generated by $\{\delta e_1, \ldots, \delta e_n\}$. Since, from Theorem 5.1, $\phi \in \mathbb{D}_{2,2}$, the sequence $(F_n, n \geq 1)$, where $F_n = E[\varphi|V_n]$, converges to φ in $\mathbb{D}_{2,2}$, hence the sequence $(M_n, n \geq 1)$, where

$$M_n = \det{}_2(I_H + \nabla^2 F_n)\exp\left\{-\mathcal{L}F_n - \frac{1}{2}|\nabla F_n|_H^2\right\},$$

converges to λ in probability. Since F_n is a 1-convex function, it follows from Theorem 6.3.1 of [22] that

$$E[g \circ (I_W + \nabla F_n)\,M_n] \leq E[g]\,,$$

for any positive, measurable function g. The first claim follows from the Fatou lemma. Since $L > 0$ almost surely, we have

$$E[g \circ T\,\Lambda] = E[g]\,, \tag{5.11}$$

for any positive, measurable g, where

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

As T is invertible, we get $\lambda \leq \Lambda$ by comparing the relations (5.10) and (5.11).
□

We can prove now the main theorem of this section:

Theorem 5.2 *Let L be given as $c^{-1}\,e^{-f}$, where $f \in \mathbb{D}_{2,1}$ is a lower bounded, finite, H-convex Wiener function and define the probability measure ν as $d\nu = Ld\mu$, where $c = E[e^{-f}]$ is the normalization constant. Let $T = I_W + \nabla\varphi$ be the optimal transportation of μ to ν in the sense of Wasserstein distance, where $\varphi \in \mathbb{D}_{2,1}$ is the 1-convex potential function. Then $\varphi \in \mathbb{D}_{2,2}$ and the Gaussian Jacobian of T is equal to $\Lambda = 1/L \circ T$ and we have the following relation:*

$$\Lambda = \det{}_2(I_H + \nabla^2\varphi)\exp\left\{-\mathcal{L}\varphi - \frac{1}{2}\,|\nabla\varphi|_H^2\right\}. \tag{5.12}$$

Proof: We have prepared everything necessary for the proof. First, we can form a sequence, denoted by $\varphi'_n, n \geq 1)$ such that each φ'_n is obtained as a convex combination from the elements of the tail sequence $(\varphi_k, k \geq n)$ and that the sequence $(\varphi'_n, n \geq 1)$ converges to φ in $\mathbb{D}_{2,2}$. Let us denote the Jacobian written with φ'_n by $\Lambda_n(\varphi'_n)$ whose explicit expression is given as

$$\Lambda_n(\varphi'_n) = \det{}_2(I + \nabla^2\varphi'_n)\exp\left\{-\mathcal{L}\varphi'_n - \frac{1}{2}|\nabla\varphi'_n|_H^2\right\}$$

Let $T'_n = I_W + \nabla\varphi'_n$ and $S'_n = I_W + \nabla\psi'_n$. Since $A \to -\log\det_2(I_H + A)$ is a convex function on the space of symmetric Hilbert-Schmidt operators which are lower bounded by $-I_H$ (cf. [3], p.63), we have

$$\begin{aligned} -\log\Lambda_n(\varphi'_n) &= -\log\det{}_2\left(I_H + \sum_i t_i\nabla^2\varphi_{n_i}\right) \\ &\quad +\sum_i t_i\mathcal{L}\varphi_{n_i} + \frac{1}{2}\left|\sum_i t_i\nabla\varphi_{n_i}\right|_H^2 \\ &\leq \sum_i -t_i\log\Lambda_{n_i}\,. \end{aligned}$$

Since $(-\log\Lambda_n, n \geq 1)$ converges to $-\log\Lambda$ in any L^p and since $(-\log\Lambda_n(\varphi'_n), n \geq 1)$ converges to $-\log\lambda$, it follows from the above inequality that

$$-\log\lambda \leq -\log\Lambda$$

almost surely, consequently $\Lambda \leq \lambda$ almost surely. It follows then from Corollary 5.1 that $\lambda = \Lambda$ almost surely and this completes the proof. □

Let us give an interesting result about the upper bound of the interpolated density whose proof makes use also the convexity results as in the proof of Theorem 5.2 :

Proposition 5.1 *Assume the hypothesis of Theorem 5.2, in particular the relation $f \geq -\alpha$. Denote by $T_t = I_W + t\nabla\varphi$, $t \in [0,1]$, then the Radon-Nikodym density L_t the measure $T_t\mu$ with respect to μ, satisfies the following inequality:*

$$L_t \leq \frac{1}{c}\exp\alpha t$$

almost surely, where $c = E[\exp -f]$.

Proof: Let g be any positive, measurable function on W, by the convexity of $t \to -\log \Lambda_t$, we have $-\log \Lambda_t \leq -t \log \Lambda$. Therefore

$$\begin{aligned} E[L_t \log L_t\, g] &= E[-\log \Lambda_t\, g \circ T_t] \\ &\leq E[-t \log \Lambda\, g \circ T_t] \\ &= E[-t(f \circ T + \log c) g \circ T_t] \\ &\leq E[(t\alpha - \log c) g \circ T_t] \\ &= E[(t\alpha - \log c) L_t\, g]\,. \end{aligned}$$

Consequently

$$L_t \log L_t \leq (t\alpha - \log c)\, L_t$$

almost surely. □

Remark 5.1 *In the finite dimensional case, if L is log-concave, it follows from [6] that $\nabla\varphi$ is a Lipschitz map (w.r.to l^1-norm). This implies that, for any $0 \leq t < 1$, the transport map T_t defined as $T_t(x) = x + t\nabla\varphi(x)$ is strongly monotone and continuous. The Theorem of Minty implies that T_t is bijective, hence from the Jacobi formula we have*

$$\int_{\mathbb{R}^n} f \circ T_t\, \Lambda_t d\mu = \int_{\mathbb{R}^n} f d\mu\,. \tag{5.13}$$

This implies in particular that

$$E[\Lambda_t] = 1\,,$$

where

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

Consequently $\mu\{L_t = 0\} = 0$ where L_t is the Radon-Nikodym derivative of $T_t\mu$ with respect to μ. Consequently $T_t\mu$ is equivalent to μ even if $T\mu$ may not. The reasoning about the bijectivity of T_t passes to the infinite dimensions without any difficulty, in fact to apply the Minty's theorem, it suffices to show the feeble-continuity in the H-direction (cf. [16] and [22]) and then to show that $\{x : T_t \circ S_t(x) = x\}$ is an H-invariant set.

5.1 An application to logarithmic Sobolev inequality

Assume now that L is H-log concave, cf. [7], then, due to the positivity improving property of the Ornsetin-Uhlenbeck semigroup, $L_n = P_{1/n}L$ is strictly positive, H-log-concave Wiener functional, hence of the form $(1/c)e^{-f}$ with $f \in \mathbb{D}_{2,1}$. Consequently the corresponding transport map T_n and $\nabla\varphi_n$ are Lipschitz maps in operator norms and their norms are both bounded by one almost surely. Denote the measure $L_n d\mu$ by $d\nu_n$. Using the logarithmic Sobolev inequality of L. Gross for the Gauss measure, we have for any nice function g (cf., also [23]):

$$
\begin{aligned}
& E_{\nu_n}\left[g^2 \log \frac{g^2}{E_{\nu_n}[g^2]}\right] \\
&= E\left[g^2\circ T_n \log \frac{g^2\circ T_n}{E[g^2\circ T_n]}\right] \\
&\leq 2E\left[|\nabla(g\circ T_n)|_H^2\right] \\
&\leq 2E\left[|\nabla g\circ T_n|_H^2 \|I_H+\nabla^2\varphi_n\|_{\text{op}}^2\right] \\
&\leq 2E\left[|\nabla g\circ T_n|_H^2\right] \\
&= 2E_{\nu_n}\left[|\nabla g|_H^2\right] \,,
\end{aligned}
$$

We can then pass to the limit as $n \to \infty$ to obtain the logarithmic Sobolev inequality for the measure ν with the minimal constant 2. Here we do not need any regularity for L, in fact, we have a similar result for the positive distributions on the Wiener space with total mass equal to one (they are Radon measures, [21]). We say that such a distribution θ is log-concave if $\theta_t = P_t\theta$ is a log-concave Wiener functional for any $t > 0$. Using the above result, we have a logarithmic Sobolev inequality for θ_t, then the inequality for θ follows by a limiting procedure, cf. [7].

5.2 Application to the exponential integrability of the Wiener functionals

The exponential integrability of the square of a random variable is an important property for applications like the large deviations. Here we give a situation which is not standart:

Theorem 5.3 *Assume that $F \in \mathbb{D}_{2,1}$ and suppose that A is a measurable subset of W which is H-convex,i.e., its indicator function is H-log-concave, such that $|\nabla F|_H \leq c$ almost surely on A. Denote by μ_A the probability measure which corresponds to the restriction of μ to A. Then we have*

$$
\mu_A\left\{|F - E_A[F]| > t\right\} \leq 2\, e^{-\frac{t^2}{2c^2}} \,,
$$

for any $t > 0$, where E_A denotes the expectation with respect to μ_A. In particular

$$
E_A\left[e^{tF^2}\right] < \infty \,,
$$

for any $t < \frac{1}{2c}$.

Proof: Let $T = I_W + \nabla\varphi$ be the transport map of Theorem 3.2. Since the set A is H-log-concave, it follows from what we have explained above that $\nabla^2\varphi$ exists as an operator valued map with $\|I_H + \nabla^2\varphi\|_{\text{op}} \leq 1$. Hence $|\nabla(F \circ T)|_H \leq |\nabla F \circ T|_H \leq c$ almost surely since $T(x) \in A$ μ a.s. Consequently, from the standart results of exponential integrability of Wiener functionals, we have

$$\mu\left\{|F \circ T - E[F \circ T]| > t\right\} \leq 2\, e^{-\frac{t^2}{2c^2}}\,,$$

and the first part of the proof is completed since $T\mu = \mu_A$. The second assertion follows easily from the first one. □

6 Itô-solutions of the Monge-Ampère equation

In the following calculations we shall take W as the classical Wiener space $W = C_0([0,1], \mathbb{R})$, $H = H^1$, i.e., the Sobolev space $W_{2,1}([0,1])$. We note that this choice does not entail any restriction of generality as indicated in [22], Chapter 2.6. Suppose we are given a positive random variables $L = \frac{1}{c}e^{-f}$ whose expectation is equal to one, c being the normalization constant. Define the measure ν as $d\nu = Ld\mu$. We shall suppose that the Wasserstein distance $d_H(\mu, \nu)$ is finite, hence the conclusions of Theorem 3.1 are valid. In order to simplify the discussion we shall assume that L is strictly positive. The transport map T can be represented as $T = I_W + \nabla\phi$ again with $\phi \in \mathbb{D}_{2,1}$. Define now

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

We have

$$\int g \circ T\ \Lambda\, d\mu = \int g\, d\mu\,,$$

for any $g \in C_b(W)$. This implies that the process $(T_t, t \in [0,1])$ defined on $[0,1] \times W$ by

$$(t, x) \to T_t(x) = x(t) + \int_0^t D_\tau\varphi(x)d\tau\,,$$

is a Wiener process under the measure $\Lambda d\mu$ with respect to its natural filtration $(\mathcal{F}_t^T, t \in [0,1])$, where $D_t\varphi$ represents the Lebesgue density of the map $t \to \nabla\varphi(x)(t) \in H$ on $[0,1]$. Since T is invertible, we have also

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

upto μ-negligeable sets. Since $\Lambda d\mu$ is equivalent to the Wiener measure, the process $(T_t, r \in [0,1])$ is a μ-semimartingale with respect to its natural filtration. It is clear that it has a decomposition of the form

$$T_t = B_t^T + A_t\,,$$

with respect to μ, where B^T is a μ-Brownian motion and A is a process of finite variation. Since we are dealing with the Brownian filtrations, $(A_t, t \in [0,1])$ should be absolutely continuous with respect to the Lebesgue measure dt of $[0,1]$. In order to calculate its density it suffices to calculate the limit

$$\lim_{h\to 0} \frac{1}{h} E\left[T_{t+h} - T_t | \mathcal{F}_t^T\right] .$$

To calculate this limit, it is enough to test it on the functions of the type $g \circ T_t$:

$$\begin{aligned} E\left[(T_{t+h} - T_t)\, g \circ T_t\right] &= E\left[(W_{t+h} - W_t)\, g \circ W_t\, L\right] \\ &= E\left[(\delta U_{[t,t+h]}) g \circ W_t\, L\right] \\ &= E\left[(U_{[t,t+h]}, \nabla(L\, g \circ W_t))_H\right] \\ &= E\left[g \circ W_t \int_t^{t+h} D_\tau L d\tau\right] , \end{aligned}$$

where $U_{[t,t+h]}$ is the element of H whose Lebesgue density is equal to the indicator function of the interval $[t, t+h]$. Hence we have

$$\begin{aligned} \lim_{h\to 0} \frac{1}{h} E\left[T_{t+h} - T_t | \mathcal{F}_t^T\right] &= -E[D_t f \circ T | \mathcal{F}_t^T] \\ &= -E_\nu[D_t f | \mathcal{F}_t] \circ T , \end{aligned}$$

$dt \times d\mu$-almost surely, where the last inequality follows from the fact that $T^{-1}(\mathcal{F}_t) = \mathcal{F}_t^T$. Hence we have proven

Proposition 6.1 *The transport process $(T_t, t \in [0,1])$ is a $(\mu, (\mathcal{F}_t^T))$-semimartingale with its canonical decomposition*

$$\begin{aligned} T_t &= B_t^T - \int_0^t E_\nu[D_\tau f | \mathcal{F}_t] \circ T \; d\tau \\ &= B_t^T - \int_0^t E[D_\tau f \circ T | \mathcal{F}_t^T] \; d\tau . \end{aligned}$$

We can give now the Itô solution of the Monge-Ampère equation:

Theorem 6.1 *Assume that $f \in \mathbb{D}_{2,1}$ be such that $c = E[\exp(-f)] < \infty$, denote by L the probability density defined by $\frac{1}{c} e^{-f}$ and by ν the probability $d\nu = L d\mu$. Assume that $d_H(\mu, \nu) < \infty$ and let $T = I_W + \nabla\varphi$ be the transport map whose properties are announced in Theorem 3.2. We have then*

$$\Lambda = \exp\left\{\int_0^1 E_\nu[D_t f | \mathcal{F}_t] \circ T dB_t^T - \frac{1}{2}\int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 \circ T\, dt\right\} . \qquad (6.14)$$

Proof: From the Itô representation formula [20], we have

$$L = \exp\left\{-\int_0^1 E_\nu[D_t f|\mathcal{F}_t]dW_t - \frac{1}{2}\int_0^1 E_\nu[D_t f|\mathcal{F}_t]^2\, dt\right\}.$$

Since the Girsanov measure for T has the density Λ given by

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

we have, using the identity $T^{-1}(\mathcal{F}_t) = \mathcal{F}_t^T$ and Proposition 6.1,

$$\begin{aligned} L\circ T &= \exp\left\{-\int_0^1 E_\nu[D_t f|\mathcal{F}_t]\circ T dT_t - \frac{1}{2}\int_0^1 E_\nu[D_t f|\mathcal{F}_t]^2\circ T\, dt\right\} \\ &= \exp\left\{-\int_0^1 E_\nu[D_t f|\mathcal{F}_t]\circ T\left(dB_t^T - E_\nu[D_t f|\mathcal{F}_t]\circ T dt\right)\right. \\ &\quad \left. -\frac{1}{2}\int_0^1 E_\nu[D_t f|\mathcal{F}_t]^2\circ T\, dt\right\}, \end{aligned}$$

which is exactly the inverse of the expression given by the relation (6.14). □

Let us give some immediate consequences of these results whose proof follows immediately from the results of this section and from Theorem 5.2 :

Corollary 6.1 *We have the following identity*

$$\begin{aligned} -\log E[e^{-f}] &= E\left[f\circ T + \frac{1}{2}\int_0^1 E_\nu[D_t f|\mathcal{F}_t]^2\circ T\, dt\right] \\ &= E\left[f\circ T + \frac{1}{2}\int_0^1 E[D_t f\circ T|\mathcal{F}_t^T]^2\, dt\right]. \end{aligned}$$

If, furthermore, f is H-convex, then we also have

$$-\log E[e^{-f}] = E\left[f\circ T - \log\det{}_2(I_H + \nabla^2\varphi) + \frac{1}{2}|\nabla\varphi|_H^2\right].$$

In particular we have the exact characterization of the Wasserstein distance between μ and ν:

$$\frac{1}{2}d_H^2(\mu,\nu) = E\left[\log\det{}_2(I_H + \nabla^2\varphi)\right] + \frac{1}{2}E_\nu\left[\int_0^1 E_\nu[D_t f|\mathcal{F}_t]^2\circ T\, dt\right].$$

7 Laplace's asymptotics and measure transportation

Let $(\Omega, \mathcal{F}, P)$ be a be a probability space and let f be a nice random variable. Then there is a well-known variational formula which says that

$$-\log \int_\Omega e^{-f} dP = \inf_{Q \in M_1(\Omega)} \left\{ R(Q|P) + \int_\Omega f dQ \right\} ,$$

where $M_1(\Omega)$ denotes the set of probability measures on $(\Omega, \mathcal{F})$ and $R(Q|P)$ is the relative entropy, i.e.,

$$R(Q|P) = \int \log \frac{dQ}{dP} dQ \,,$$

whenever this quantity has a sense, otherwise it is defined as to be identically infinity. Besides, the above infimum is attained at the probability whose Radon-Nikodym derivative is given by

$$\frac{dQ}{dP} = e^{-f} \frac{1}{\int e^{-f} dP} \,.$$

Assume now that we are working in an abstract Wiener space (W, μ, H) and let the probability P be the Wiener measure μ. The above problem can be transformed to the minimization problem

$$\inf \left\{ E[K \log K] + E[f\, K] : \; K \in \mathbb{L}^1_{1,+} \right\} \,.$$

Since the above infimum will be calculated necessarily between the random variables $K \in \mathbb{L}^1_{1,+}$ with $E[K \log K] < \infty$, for any such K, from Theorem 3.2, there exists a unique 1-convex $F \in \mathbb{D}_{2,1}$, such that

$$\frac{d(I + \nabla F)\mu}{d\mu} = K \,.$$

Therefore

$$E[K \log K] + E[f\, K] = E\left[-\log \Lambda_F + f \circ (I + \nabla F) \right] ,$$

where

$$\Lambda_F = \frac{1}{K \circ (I + \nabla F)} \,.$$

In the finite dimensional case, without any further regularity assumption on f, we can write Λ_F as (cf., [14, 24] and [9]):

$$\Lambda_F = \det\nolimits_2 (I + \nabla_a^2 F) \exp \left\{ -\mathcal{L}_a F - \frac{1}{2} |\nabla F|^2 \right\} ,$$

where the subscript refers to the Alexandrov version of the corresponding operations. Let us show that the infimum above is attained and calculate it: let $T = I + \nabla\varphi$ be the optimal transport of the measure $d\mu$ to $Ld\mu$, where L is defined by

$$L = e^{-f}\frac{1}{\int e^{-f}d\mu}\,.$$

Then we have

$$\begin{aligned}\Lambda_\varphi &= \frac{1}{L\circ T}\\ &= e^{f\circ T}E\left[e^{-f}\right]\\ &= \det{}_2(I+\nabla_a^2\varphi)\exp\left\{-\mathcal{L}_a\varphi - \frac{1}{2}|\nabla\varphi|_H^2\right\}\,.\end{aligned}$$

Solving $E[e^{-f}]$ from above, we get

$$-\log E[e^{-f}] = f\circ T - \log\det{}_2(I+\nabla_a^2\varphi) + \mathcal{L}_a\varphi + \frac{1}{2}|\nabla\varphi|^2\,. \tag{7.15}$$

Taking the expectation of the both sides of (7.15), we obtain

$$-\log E[e^{-f}] = E\left[f\circ T - \log\det{}_2(I+\nabla_a^2\varphi) + \mathcal{L}_a\varphi\right] + \frac{1}{2}d_H^2(\mu, L\cdot\mu)\,,$$

where $L\cdot\mu$ denotes the measure whose density with respect to μ is L. Assume now that f is a convex function, i.e., that L is a log-concave function. Then from [6], the potential function φ is in $\mathbb{D}_{2,2}$, in fact $\nabla\varphi$ is 1-Lipschitz, i.e.,

$$|\nabla\varphi(x+h) - \nabla\varphi(x)|_H \le |h|_H\,,$$

for any $x \in W$, $h \in H$. Hence in the above expressions we can replace $\nabla_a^2\varphi$ and $\mathcal{L}_a\varphi$ with $\nabla^2\varphi$ and $\mathcal{L}\varphi$ respectively and since $E[\mathcal{L}\varphi] = 0$, we get

$$-\log E[e^{-f}] = E\left[f\circ T - \log\det{}_2(I+\nabla^2\varphi)\right] + \frac{1}{2}d_H^2(\mu, L\cdot\mu)\,. \tag{7.16}$$

Remark 7.1 *In this latter case we have*

$$-\log E[e^{-f}] = f\circ T + \frac{1}{2}|\nabla\varphi|^2 + \mathcal{L}\varphi - \log\det{}_2(I+\nabla^2\varphi)\,. \tag{7.17}$$

Since the left hand side of this equality is constant, taking the (distributional) derivative of this relation, we see that the expression

$$\mathcal{L}\varphi - \log\det{}_2(I+\nabla^2\varphi)$$

has the same regularity as the term $f\circ T + 1/2|\nabla\varphi|^2$.

Let now

$$a(n) = \int_{\mathbb{R}^d} \exp\left\{-nf(x/\sqrt{n})\right\} d\mu(x)\,,$$

for $n \in \mathbb{N}$. It follows from Laplace's method, cf.,for instance [17], that

$$\lim_{n\to\infty} \left[-\frac{1}{n}\log a(n)\right] = \inf_{x\in\mathbb{R}^d}\left\{\frac{1}{2}|x|^2 + f(x)\right\}.$$

Let now ν_n be the measure defined by

$$d\nu_n(x) = \frac{1}{a(n)}\exp\left\{-nf(x/\sqrt{n})\right\} d\mu(x)\,.$$

Denote by $T_n = I_{\mathbb{R}^d} + \nabla\varphi_n$ the transport map such that $T_n\mu = \nu_n$. It follows from the relation (7.16) that

$$\begin{aligned}
\lim_{n\to\infty} \quad & \left\{\int_{\mathbb{R}^d}\left[f\left(\frac{T_n(x)}{\sqrt{n}}\right) - \frac{1}{n}\log\det{}_2(I+\nabla\varphi_n(x))\right]\mu(dx)\right. \\
& \left. + \frac{1}{2n}d_H^2(T_n\mu,\mu)\right\} \\
= \quad & \inf_{x\in\mathbb{R}^d}\left\{\frac{1}{2}|x|^2 + f(x)\right\}.
\end{aligned}$$

Note that the right hand side of the above equality is nothing but the negative of the dual potential function at $y = 0$, i.e., if we define g as

$$g(y) = -\inf_{x\in\mathbb{R}^d}\left\{\frac{1}{2}|x-y|^2 + f(x)\right\},$$

then

$$\lim_{n\to\infty}\left[-\frac{1}{n}\log a(n)\right] = -g(0)\,.$$

Let us now pass to the infinite dimensional situation: assume that f satisfies the hypothesis of Theorem 5.2, then with the notations of this theorem, using exactly the same arguments as above we conclude that

$$-\log E[e^{-f}] = E\left[f\circ T - D_a(\varphi) + \mathcal{L}_a\varphi\right] + \frac{1}{2}d_H^2(L\cdot\mu,\mu)\,. \tag{7.18}$$

Since, from the proof of Theorem 5.2, we have $\mathcal{L}\varphi - D(\varphi) = \mathcal{L}_a\varphi - D_a(\varphi)$, we can write the relation (7.18) also as

$$-\log E[e^{-f}] = E\left[f\circ T\right] - \langle D(\varphi), 1\rangle + \frac{1}{2}d_H^2(L\cdot\mu,\mu)\,, \tag{7.19}$$

where the brackets correspond to the duality $(\mathbb{D},\mathbb{D}')$ (or to $(\mathbb{D}_{2,1},\mathbb{D}_{2,-1})$).

Remark 7.2 *The following relation is proven in [5] in the frame of the classical Wiener space, equipped with its usual filtration,*

$$-\log E\left[e^{-f}\right] = \inf_u E\left[\frac{1}{2}|u|_H^2 + f \circ (I_W + u)\right],$$

where $u = \int_0^\cdot \dot{u}_s ds$, *and* $\dot{u}$ *runs in the set of adapted (progressively measurable) processes. Again if* f *is reasonably regular, then there exists a minimizing process, say* u_m *which satisfies the following equation:*

$$\dot{u}_m(t) + E\left[(D_t f) \circ (I_W + u_m)|\mathcal{F}_t\right] = 0, \tag{7.20}$$

$dt \times d\mu$*-almost surely. Using the fixed point theorem, it is easy to show that the equation (7.20) has a unique solution in* $\mathbb{D}_{2,0}(H)$ *if* ∇f *has a derivative bounded by some* $c < 1$*,*

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] E.F. Beckenbach and R. Bellman: *Inequalities*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 30. Springer-Verlag, 1983.

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

[5] M. Boué and P. Dupuis: "A variational representation for certain functionals of Brownian motion". The Annals of Probability, **26**, No. 4, 1641-1659, 1998.

[6] L. A. Caffarelli: "Monotonicity properties of optimal transportation and the FKG and related inequalities". Commun. Math. Phys.**214**, 547-563, 2000.

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

[8] 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.

[9] D. Feyel and A. S. Üstünel: "Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space". Preprint, to appear in Prob. Theory and Related Fields.

[10] D. Feyel and A. S. Üstünel: "Monge-Kantorovitch measure transportation, Monge-Ampère equation and the Itô calculus". To appear in Advanced Studies in Pure Mathematics, Vol. **41**. Mathematical Society of Japan.

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

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

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

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

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

[16] J. Schwartz: *Nonlinear Functional Analysis*. Gordon and Breach, 1969.

[17] B. Simon: *Functional Integration and Quantum Physics*. Academic Press, 1979.

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

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

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

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

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

[23] A. S. Üstünel: "Damped logarithmic Sobolev inequality on the Wiener space". Stochastic Analysis and Related Topics VII, The Silivri Workshop, p. 245-249. Progress in Probability, Vol. **48**. Birkhäuser, 2001.

[24] C. Villani: *Topics in Optimal Transportation*. Graduate Series in Math., **58**. Amer. Math. Soc., 2003.

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

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