



Wasserstein geometry of porous medium equation [☆]

Asuka Takatsu ^{a,b,c,*}

^a Graduate School of Mathematics, Nagoya University, Nagoya 464-8602, Japan

^b Institut des Hautes Études Scientifiques, Bures-sur-Yvette 91440, France

^c Max-Planck-Intitut für Mathematik, Vivatsgasse 7, 53111 Bonn, Germany

Received 18 February 2010; received in revised form 23 August 2011; accepted 11 October 2011

Available online 19 October 2011

Abstract

We study the porous medium equation with emphasis on q -Gaussian measures, which are generalizations of Gaussian measures by using power-law distribution. On the space of q -Gaussian measures, the porous medium equation is reduced to an ordinary differential equation for covariance matrix. We introduce a set of inequalities among functionals which gauge the difference between pairs of probability measures and are useful in the analysis of the porous medium equation. We show that any q -Gaussian measure provides a nontrivial pair attaining equality in these inequalities.

© 2011 Elsevier Masson SAS. All rights reserved.

MSC: 60D05; 46E27

Keywords: q -Gaussian measure; Porous medium equation; Functional inequality

1. Introduction

A q -Gaussian measure is one of power-law distributions on \mathbb{R}^d , which is described by mean, covariance matrix parameters and the q -exponential function given by

$$\exp_q(t) := [1 + (1 - q)t]_+^{\frac{1}{1-q}} \quad \text{for } q \in Q_d := (0, 1) \cup \left(1, \frac{d+4}{d+2}\right),$$

where we put $[x]_+ := \max\{x, 0\}$ and by convention $0^a := \infty$ for any negative number a (see (3.1) for the precise definition of q -Gaussian measures). The q -Gaussian measure is considered as an approximation of the Gaussian measure since the q -exponential function recovers the usual exponential function in the limit $q \rightarrow 1$. This paper aims to demonstrate that the q -Gaussian measure inherits some features of the Gaussian measure. To do this, let us start by recalling properties of the Gauss measure.

For any $v \in \mathbb{R}^d$ and $V \in \text{Sym}^+(d, \mathbb{R})$, which is the set of symmetric positive definite matrices of size d , the Gaussian measure $N(v, V)$ with mean v and covariance matrix V is given by

[☆] This work was partially supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists and by JSPS–IHÉS (EPDI) Fellowship.

* Correspondence to: Graduate School of Mathematics, Nagoya University, Nagoya 464-8602, Japan.
E-mail address: takatsu@math.nagoya-u.ac.jp.

$$N(v, V) := \exp\left(-\frac{1}{2}\langle x - v, V^{-1}(x - v) \rangle\right) (\det(2\pi V))^{-\frac{1}{2}} \mathcal{L}^d,$$

where \mathcal{L}^d stands for the Lebesgue measure on \mathbb{R}^d . For example, the density of $N(0, 2tI_d)$, where I_d is the identity matrix of size d and $t > 0$, is the heat kernel which is a self-similar solution to the heat equation

$$\frac{\partial}{\partial t} \rho = \Delta \rho.$$

As well as the heat kernel, a solution to the heat equation with an initial data being a Gaussian density remains Gaussian densities for all future time, that is, the space of Gaussian measures is stable under the heat equation. On the space of Gaussian measures, the heat equation is reduced to an ordinary differential equation for covariance matrix.

It simplifies not only the analysis of the heat equation but also the analysis of the Wasserstein space to restrict arguments to the space of Gaussian measures. The *Wasserstein space* is the space \mathcal{P}_2 of Borel probability measures on \mathbb{R}^d having finite second moments equipped with a distance function W_2 defined by

$$W_2(\mu, \nu) := \inf \left\{ \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\pi(x, y) \right)^{\frac{1}{2}} \mid \pi: \text{a transport plan of } \mu \text{ and } \nu \right\},$$

where a *transport plan* of μ and ν is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ such that $\pi[B \times \mathbb{R}^d] = \mu[B]$ and $\pi[\mathbb{R}^d \times B] = \nu[B]$ for any Borel set $B \subset \mathbb{R}^d$. A transport plan is said to be *optimal* if it attains the infimum above, which always exists (see [24, Chapter 4]). Though the explicit expression of an optimal transport plan is not usually obtained, the explicit expression of an optimal transport plan of a pair of Gaussian measures is known, which guarantees the convexity of the space of Gaussian measures in Wasserstein geometry. The restriction to the space makes it possible to analyze Wasserstein geometry in detail (see [17]).

We furthermore focus on the fact that the Gaussian measure satisfies a set of inequalities among the relative entropy H , the Fisher information I and the Wasserstein distance function W_2 , all of which gauge differences between pairs of probability measures and are defined by

$$H(\mu|\nu) := \begin{cases} \int_{\mathbb{R}^d} \frac{d\mu}{d\nu} \ln \frac{d\mu}{d\nu} d\nu & \text{if } \mu \text{ is absolutely continuous with respect to } \nu, \\ +\infty & \text{otherwise,} \end{cases}$$

$$I(\mu|\nu) := \begin{cases} \int_{\mathbb{R}^d} |\nabla \ln \frac{d\mu}{d\nu}|^2 d\mu & \text{if } \mu \text{ is absolutely continuous with respect to } \nu, \\ +\infty & \text{otherwise.} \end{cases}$$

We say that a probability measure ν *satisfies the logarithmic Sobolev inequality* with constant λ , in short $LS(\lambda)$ if we have

$$H(\mu|\nu) \leq \frac{1}{2\lambda} I(\mu|\nu)$$

for all absolutely continuous probability measure μ with respect to ν . Similarly, the probability measure ν is said to *satisfy the Talagrand inequality* with constant λ , in short $T(\lambda)$ (resp. the *HWI inequality* with constant λ , in short $HWI(\lambda)$) if we have

$$W_2(\mu, \nu) \leq \sqrt{\frac{2}{\lambda} H(\mu|\nu)} \quad \left(\text{resp. } H(\mu|\nu) \leq W_2(\mu, \nu) \sqrt{I(\mu|\nu)} - \frac{\lambda}{2} W_2(\mu, \nu)^2 \right)$$

for all absolutely continuous probability measure μ with respect to ν . Any Gaussian measure satisfies $LS(\lambda)$, $T(\lambda)$ and $HWI(\lambda)$, where λ^{-1} is the largest eigenvalues of the covariance matrix, and moreover provides a nontrivial pair attaining equality in these inequalities. Note that criteria for a probability measure to satisfy $LS(\lambda)$, $T(\lambda)$ and $HWI(\lambda)$ are known, for example, $LS(\lambda)$ with some convexity condition for ν implies $T(\lambda)$ (see [15] for the details). The key of the proof that $LS(\lambda)$ implies $T(\lambda)$ is to analyze asymptotic behaviors of the Fokker–Planck equation of the form

$$\frac{\partial}{\partial t} \rho = \Delta \rho + \text{div}(\rho \nabla \Psi),$$

where Ψ is a function on \mathbb{R}^d , on the contrary, these inequalities are applied to analyze asymptotic behaviors of the Fokker–Planck equation. In particular, Otto [14] investigated the asymptotic behaviors in the case that $\Psi(x) = |x|^2/4$ using Wasserstein gradient structure, where the scaling

$$\rho(t, x) = t^{-\frac{d}{2}} \hat{\rho}(\ln t, t^{-\frac{1}{2}}x)$$

provides a one-to-one correspondence between solutions ρ to the heat equation and the solutions $\hat{\rho}$ to the Fokker–Planck equation.

We show that such features are inherited to the space of q -Gaussian measures after some preliminaries on the q -exponential function and the Wasserstein geometry in Section 2. Section 3 is devoted to the convexity of the space of q -Gaussian measures in the Wasserstein space (Theorem A). In Section 4, we consider the features of q -Gaussian measures as solutions to the porous medium equation

$$\frac{\partial}{\partial t} \rho = \Delta(\rho^{2-q}). \tag{1.1}$$

Since the stability of the space of q -Gaussian measures has been already shown by Ohara and Wada [12], it is possible to restrict the porous medium equation to the space of q -Gaussian measures. We introduce the ordinary differential equation for covariance matrix obtained by restricting the porous medium equation to the space of q -Gaussian measures (Theorem B). Section 5 is concerned with generalizations of the logarithmic Sobolev inequality, the Talagrand inequality and the HWI inequality, which play crucial roles in analyzing the asymptotic behaviors of the nonlinear Fokker–Planck equation

$$\frac{\partial}{\partial t} \rho = \Delta(\rho^{2-q}) + \operatorname{div}(\rho \nabla \Psi).$$

We give criteria for a probability measure to satisfy these inequalities and show that any q -Gaussian measure provides a nontrivial pair attaining equality in these inequalities if $q > 1$ (Corollaries C, D). We finally prove that a variant of the logarithmic Sobolev inequality implies a variant of the Talagrand inequality using the nonlinear Fokker–Planck equation (Theorem E).

We refer to the preceding results in the literature. Generalizations of the logarithmic Sobolev inequality, the Talagrand inequality and the HWI inequality have been studied by several authors. Carrillo, Jüngel, Markowich, Toscani and Unterreiter [6] studied these functional inequalities for parabolic systems using entropy dissipation methods. Carrillo, McCann and Villani [7] also investigated these functional inequalities and they also estimated in [8] the contraction rate of nonlinear evolution equations in the Wasserstein distance function. Agueh, Ghoussoub and Kang [1] and Cordero-Erausquin, Gangbo and Houdré [9] introduced variants of the relative entropy and the Fisher information, through modifying the Boltzmann entropy and the quadratic transport cost. However, none of them referred to the q -Gaussian measures and our result is the first one concerning the importance of the q -Gaussian measures in the porous medium equation. See also [13], where Ohta and the author investigated the nonlinear Fokker–Planck equation on a (weighted) Riemannian manifold using the Wasserstein gradient structure and the q -exponential function.

2. Preliminaries

2.1. q -exponential function and q -logarithmic function

We first summarize the q -calculus, see [21] for further discussion. Take $q \in Q_d$ and fix it. We define the q -logarithmic function \ln_q by

$$\ln_q(t) := \frac{t^{1-q} - 1}{1 - q}$$

for $t > 0$. Since the function \ln_q is monotone increasing, there exists its inverse function on the image. This inverse function is called the q -exponential function \exp_q and is naturally extended to all of \mathbb{R} as

$$\exp_q(t) := [1 + (1 - q)t]_+^{\frac{1}{1-q}}.$$

Note that the functions \ln_q and \exp_q recover the usual logarithmic function and the usual exponential function as q tends to 1, respectively.

Let us define functionals on the space \mathcal{P}^{ac} of absolutely continuous probability measures with respect to the Lebesgue measure using the q -logarithmic function. The Tsallis entropy E_q is defined by

$$E_q(\mu) := - \int_{\mathbb{R}^d} f^q \ln_q(f) d\mathcal{L}^d = - \int_{\mathbb{R}^d} \frac{f^q - f}{q - 1} d\mathcal{L}^d$$

for $\mu = f\mathcal{L}^d \in \mathcal{P}^{\text{ac}}$, which is regarded as an approximation of the Boltzmann entropy E since we have

$$E_q(\mu) \xrightarrow{q \rightarrow 1} E(\mu) = - \int_{\mathbb{R}^d} f \ln(f) d\mathcal{L}^d.$$

We also define the q -relative entropy H_q and the q -Fisher information by

$$H_q(\mu|\nu) := \frac{1}{2-q} \int_{\mathbb{R}^d} [f \ln_q(f) - g \ln_q(g) - (2-q) \ln_q(g)(f-g)] d\mathcal{L}^d,$$

$$I_q(\mu|\nu) := \int_{\mathbb{R}^d} |\nabla[\ln_q(f) - \ln_q(g)]|^2 d\mu$$

for $\mu = f\mathcal{L}^d, \nu = g\mathcal{L}^d \in \mathcal{P}^{\text{ac}}$, both of which are non-negative and gauge the difference between pairs of probability measures. Note that $\lim_{q \rightarrow 1} H_q(\mu|\nu) = H(\mu|\nu)$ and $\lim_{q \rightarrow 1} I_q(\mu|\nu) = I(\mu|\nu)$ hold. The square root of the q -relative entropy H_q is considered as a generalization of the distance function in the context of information geometry since this satisfies a generalized Pythagorean relation (see [2,3] for information geometry). We refer to the generalized Pythagorean relation in Remark 3.2. It will be demonstrated in (5.15) that $-I_q(\mu|\nu)$ is the first variation of $H_q(\cdot|\nu)$ at μ .

2.2. Wasserstein geometry

We briefly recall some results on the Wasserstein space. See [23,24] and references therein for the details and more information on Wasserstein geometry.

As mentioned in the introduction, an optimal transport of any pair in \mathcal{P}_2 always exists. Though the explicit expression of an optimal transport is usually not obtained, it has been known since the work of Brenier [5] that an optimal transport is characterized by a push forward measure of the gradient of a convex function. A *push forward measure* of a probability measure μ by a map T , denoted by $T_{\#}\mu$, is defined by $T_{\#}\mu[B] := \mu[T^{-1}(B)]$ for all Borel sets $B \subset \mathbb{R}^d$. Given two probability measures $\mu = f\mathcal{L}^d, \nu = g\mathcal{L}^d$ and a differentiable map T on $\mathbb{R}^d, \nu = T_{\#}\mu$ is equivalent to that

$$f = g(T) \det(dT) \tag{2.1}$$

holds for μ -almost everywhere, where dT is the total differential of T . We denote by id the identity map on \mathbb{R}^d .

Theorem 2.1. (See [5].) *Let $\mu, \nu \in \mathcal{P}_2$ be such that μ does not give mass to sets of Hausdorff dimension at most $(d - 1)$. Then there exists a convex function such that $\nabla\varphi$ pushes μ forward to ν and $[\text{id} \times \nabla\varphi]_{\#}\mu$ is a unique optimal transport plan of them. Moreover, $\{[(1 - t)\text{id} + t\nabla\varphi]_{\#}\mu\}_{t \in [0,1]}$ is a unique geodesic from μ to ν in the Wasserstein space.*

The converse situation also holds true, that is, if a support of a transport plan is almost contained in the subdifferential of a convex function, then the transport plan is optimal. This is known as Knott–Smith optimality criterion.

Theorem 2.2. (See [23, Theorem 2.16].) *Let φ be a proper lower semicontinuous convex function on \mathbb{R}^d . For $\mu, \nu \in \mathcal{P}_2$, let π be a transport plan of them such that*

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \left[\varphi(x) + \sup_{z \in \mathbb{R}^d} (\langle y, z \rangle - \varphi(z)) - \langle x, y \rangle \right] d\pi(x, y) \leq \varepsilon.$$

Then we have

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\pi(x, y) \leq W_2(\mu, \nu)^2 + 2\varepsilon.$$

We finally give the explicit expression of the optimal transport between a pair of Gaussian measures, which helps us to understand Wasserstein geometry (see [17] and references therein for more details). Given any $X \in \text{Sym}^+(d, \mathbb{R})$, we define a symmetric positive definite matrix $X^{1/2} = \sqrt{X}$ such that $X^{1/2} \cdot X^{1/2} = X$.

For any pair of Gaussian measures $N(v, V)$ and $N(u, U)$, we define the symmetric positive definite matrix T and the associated function \mathcal{T} by

$$T := U^{\frac{1}{2}}(U^{\frac{1}{2}}VU^{\frac{1}{2}})^{-\frac{1}{2}}U^{\frac{1}{2}}, \quad \mathcal{T}(x) := \frac{1}{2}\langle x - v, T(x - v) \rangle + \langle x, u \rangle, \tag{2.2}$$

which provides an optimal transport plan of $N(v, V)$ and $N(u, U)$. In other words, $[\text{id} \times \nabla \mathcal{T}]_{\#}N(v, V)$ is the optimal transport plan of them and then the Wasserstein distance between them is given by

$$W_2(N(v, V), N(u, U))^2 = |v - u|^2 + \text{tr } V + \text{tr } U - 2 \text{tr} \sqrt{U^{\frac{1}{2}}VU^{\frac{1}{2}}}.$$

A unique geodesic from $N(v, V)$ to $N(u, U)$ is given by $\{N(w_t, W_t)\}_{t \in [0, 1]}$, where the time-dependent vector w_t and the time-dependent matrix W_t are defined by

$$w_t := (1 - t)v + tu, \quad W_t := [(1 - t)I_d + tT]V[(1 - t)I_d + tT]. \tag{2.3}$$

3. q -Gaussian measure

Let us summarize the definition of q -Gaussian measures and then discuss Wasserstein geometry of the space of q -Gaussian measures. Background on q -Gaussian measures is found in [20] and [21].

A probability measure $N_q(v, V)$ is called the q -Gaussian measure with mean v and covariance matrix V if it maximizes the Tsallis entropy E_q among $\mu \in \mathcal{P}^{\text{ac}}$ with mean v and covariance matrix V . It is known that the q -Gaussian measure $N_q(v, V)$ is given by

$$N_q(v, V) = C_0(\det V)^{-\frac{1}{2}} \exp_q \left[-\frac{1}{2}C_1 \langle x - v, V^{-1}(x - v) \rangle \right] \mathcal{L}^d, \tag{3.1}$$

where C_0 and C_1 are the positive constants given by

$$C_0 = C_0(q, d) := \begin{cases} \frac{\Gamma(\frac{2-q}{1-q} + \frac{d}{2})}{\Gamma(\frac{2-q}{1-q})} \left(\frac{(1-q)C_1}{2\pi} \right)^{\frac{d}{2}} & \text{if } 0 < q < 1, \\ \frac{\Gamma(\frac{1}{q-1})}{\Gamma(\frac{1}{q-1} - \frac{d}{2})} \left(\frac{(q-1)C_1}{2\pi} \right)^{\frac{d}{2}} & \text{if } 1 < q < \frac{d+4}{d+2}, \end{cases}$$

$$C_1 = C_1(q, d) := \frac{2}{2 + (d + 2)(1 - q)}$$

and $\Gamma(\cdot)$ is the Γ -function (see [22]). In some cases, the q -Gaussian measure $N_q^*(v, V)$ is obtained as a maximizer of the Tsallis entropy E_q under the q -mean v constraint and the q -covariance V constraint, that is, $N_q^*(v, V)$ maximizes the Tsallis entropy E_q among $\mu = f \mathcal{L}^d \in \mathcal{P}^{\text{ac}}$ satisfying

$$\begin{cases} Q(f) = \int_{\mathbb{R}^d} f^q d\mathcal{L}^d, \\ \int_{\mathbb{R}^d} x f(x)^q d\mathcal{L}^d(x) = Q(f)v, \\ \int_{\mathbb{R}^d} (x - v)^{\top}(x - v) f(x)^q d\mathcal{L}^d(x) = Q(f)V, \end{cases}$$

where vectors in \mathbb{R}^d are column and ${}^{\top}x$ stands for the transpose of x . A relation between $N_q(v, V)$ and $N_q^*(v, V)$ is given by

$$N_q^*(v, V) = N_q \left(v, \frac{2 + d(1 - q)}{2 + (d + 2)(1 - q)} V \right).$$

We use the usual mean and covariance constraint condition throughout this paper. In this case, the q -Gaussian measure is well defined for $q \in Q_d$ and the q -Gaussian measure $N_q(v, V)$ recovers the Gaussian measure $N(v, V)$ as q tends to 1. We denote by $n_q(v, V)$ the density of $N_q(v, V)$ with respect to the Lebesgue measure.

The space of q -Gaussian measures is convex in the Wasserstein space.

Theorem A. *For any $q \in Q_d$, the space of q -Gaussian measures is convex and isometric to the space of Gaussian measures with respect to Wasserstein geometry.*

Proof. For the function T given in (2.2), we have

$$n_q(v, V) = n_q(u, U)(\nabla T)(\det \text{Hess } T)$$

for $N_q(v, V)$ -almost everywhere, which implies $[\nabla T]_{\#} N_q(v, V) = N_q(u, U)$ by (2.1). Theorem 2.2 with the convexity of T ensures the optimality of $[\text{id} \times \nabla T]_{\#} N_q(v, V)$ and we have

$$\begin{aligned} W_2(N_q(v, V), N_q(u, U))^2 &= |v - u|^2 + \text{tr } V + \text{tr } U - 2 \text{tr} \sqrt{U^{\frac{1}{2}} V U^{\frac{1}{2}}} \\ &= W_2(N(v, V), N(u, U))^2. \end{aligned}$$

Hence the map from the space of Gaussian measures to the space of q -Gaussian measures sending $N(v, V)$ to $N_q(v, V)$ is an isometry with respect to W_2 .

Moreover, for the time-dependent vector w_t and the time-dependent matrix W_t given in (2.3), $\{N_q(w_t, W_t)\}_{t \in [0,1]}$ is a unique geodesic from $N_q(v, V)$ to $N_q(u, U)$, which shows the convexity of the space of q -Gaussian measures in the Wasserstein space. \square

Remark 3.1. The convexity of the space of q -Gaussian measures is due to the characterization by the mean and covariance matrix parameter rather than the q -exponential function, which suggests the existence of other convex spaces (see [19]).

Remark 3.2. We briefly explain that the square root of the q -relative entropy satisfies a generalized Pythagorean relation. Given $\mu \in \mathcal{P}^{\text{ac}}$, let $N_q(v, V)$ be a minimizing q -Gaussian measure for the variational problem

$$\min\{H_q(\mu | N_q(u, U)) \mid (u, U) \in \mathbb{R}^d \times \text{Sym}^+(d, \mathbb{R})\}.$$

Then the following Pythagorean relation

$$H_q(\mu | N_q(u, U)) = H_q(\mu | N_q(v, V)) + H_q(N_q(v, V) | N_q(u, U))$$

holds for any q -Gaussian measure $N_q(u, U)$. For the proof, see [12, Proposition 3].

4. q -Gaussian measures as solutions to porous medium equation

It is known that the porous medium equation (1.1) allows for a self-similar solution of the form

$$\rho_q(x, t) := [At^{-d\alpha(1-q)} - B|x|^2t^{-1}]_+^{\frac{1}{1-q}} = [A - B|x|^2t^{-2\alpha}]_+^{\frac{1}{1-q}} t^{-d\alpha},$$

where the constants α and B are given by

$$\alpha = \alpha(q, d) := \frac{1}{d(1-q) + 2}, \quad B = B(q, d) := \frac{(1-q)\alpha}{2(2-q)}.$$

The other constant $A = A(q, d)$ is defined by the total mass of the solution and we normalized it such that

$$\int_{\mathbb{R}^d} \rho_q(x, t) d\mathcal{L}^d = 1.$$

To be precise, A is given by

$$A := C_0^{2\alpha(1-q)} \left[\frac{\alpha}{(2-q)C_1} \right]^{d\alpha(1-q)}.$$

The solution ρ_q was discovered by Barenblatt [4] and Pattle [16] and is called the *Barenblatt–Pattle solution*. It is easy to check the relation

$$\rho_q(x, t) = n_q(0, Ct^{2\alpha}I_d)(x),$$

where the constant $C = C(q, d)$ is given by

$$C := \frac{(2-q)C_1}{\alpha} A.$$

As well as the Barenblatt–Pattle solution, it was proved by Ohara and Wada [12, Proposition 5] that a solution to the porous medium solution with an initial data being a q -Gaussian density remains q -Gaussian densities for all future time. This fact implies that a solution to the porous medium equation on the space of q -Gaussian measures can be explicitly solved [12, Remark 2]. To do this, we introduce the map Θ on $\text{Sym}^+(d, \mathbb{R})$ such that

$$\Theta(V) := (\det V)^{-\alpha(1-q)} V.$$

Note that $\rho_q(x, t) = n_q(0, C\Theta(tI_d))(x)$ holds.

Theorem B. For any $q \in Q_d$ and $V \in \text{Sym}^+(d, \mathbb{R})$, we set the time-dependent matrix V_t as

$$\Theta(V_t) = \Theta(V) + \sigma(t)I_d, \quad \frac{d}{dt}\sigma(t) = 2\alpha(\det \Theta(V_t))^{-\frac{1-q}{2}}.$$

Then $n_q(v, C\Theta(V_t))$ is a solution to the porous medium equation (1.1).

Remark 4.1. The assertion also holds true for $q = 1$.

Proof of Theorem B. For simplicity, we use the following notations

$$|x|_V^2 := \langle x, V^{-1}x \rangle, \quad \Theta_t := \Theta(V_t), \quad F(t, x) := [A - B|x - v|_{\Theta_t}^2]_+.$$

Then we have $\rho(t, x) := F^{1/(1-q)}(\det \Theta_t)^{-1/2} = n_q(v, C\Theta(V_t))(x)$ and calculate

$$\Delta(\rho^{2-q}(t, x)) = \alpha F^{\frac{q}{1-q}}(t, x)(\det \Theta_t)^{-\frac{2-q}{2}} \left(\frac{2B}{1-q}|x - v|_{\Theta_t}^2 - F(t, x)\text{tr}(\Theta_t^{-1}) \right).$$

Combining well-known equations for a time-dependent invertible matrix X_t

$$\frac{d}{dt}(X_t^{-1}) = -X_t^{-1} \left(\frac{d}{dt} X_t \right) X_t^{-1}, \quad \frac{d}{dt} \det X_t = (\det X_t) \text{tr} \left(X_t^{-1} \frac{d}{dt} X_t \right)$$

with the assumption

$$\frac{d}{dt} \Theta_t = 2\alpha(\det \Theta_t)^{-\frac{1-q}{2}} I_d,$$

we obtain

$$\frac{\partial}{\partial t}(\rho(t, x)) = \alpha F^{\frac{q}{1-q}}(t, x)(\det \Theta_t)^{-\frac{2-q}{2}} \left(\frac{2B}{1-q}|x - v|_{\Theta_t}^2 - F(t, x)\text{tr}(\Theta_t^{-1}) \right).$$

We thus have

$$\frac{\partial}{\partial t} \rho = \Delta(\rho^{2-q}),$$

proving that $\rho(t, x) = n_q(c, C\Theta(V_t))(x)$ is a solution to (1.1). \square

Remark 4.2. It is known that the scaling

$$\rho(t, x) = t^{-d\alpha} \hat{\rho}(\ln t, t^{-\alpha} x)$$

provides a one-to-one correspondence between solutions ρ to the porous medium equation and solutions $\hat{\rho}$ to the nonlinear Fokker–Planck equation of the form

$$\frac{\partial}{\partial t} \hat{\rho} = \Delta(\hat{\rho}^{2-q}) + \left[\hat{\rho} \left(\nabla \frac{\alpha}{2} |x|^2 \right) \right].$$

In particular, the Barenblatt–Pattle solution corresponds to the stationary solution $n_q(0, CI_d)$ to the nonlinear Fokker–Planck equation (for instance, see [14]). Since the solutions obtained by Theorem B contain nonself-similar solutions, we obtain nonequilibrium solutions to the nonlinear Fokker–Planck equation. Such solutions help us to understand the asymptotic behavior of solutions to the nonlinear Fokker–Planck equation. As an example, the author [18] demonstrated by using elemental calculations that $N_q(0, CI_d)$ satisfies the q -logarithmic Sobolev inequality, the q -Talagrand inequality and the q -HWI inequality for any q -Gaussian measures, which are defined in Section 5 and are key ingredients to analyze the asymptotic behavior of solutions to the nonlinear Fokker–Planck equation.

5. Functional inequalities

In this section, we study relations among the functionals H_q, I_q and W_2 in the form of inequality and discuss the importance of q -Gaussian measures in the inequalities. We deal with the three inequalities, called the q -logarithmic Sobolev inequality, q -Talagrand inequality and q -HWI inequality, of the form

$$\begin{aligned} H_q(\mu|v) &\leq \frac{1}{2\lambda} I_q(\mu|v), \\ W_2(\mu, v) &\leq \sqrt{\frac{2}{\lambda} H_q(\mu|v)}, \\ H_q(\mu|v) &\leq \sqrt{I_q(\mu|v) W_2(\mu, v)} - \frac{\lambda}{2} W_2(\mu, v)^2. \end{aligned}$$

Though criteria for a probability measure to satisfy the three inequalities have been already provided even in a more general setting (for instance see [1,7,9]), we show criteria to emphasize the importance of the q -Gaussian measure. We use the same symbol ∇ for the distributional gradient.

Lemma 5.1. *Given $\mu_1 = f_1 \mathcal{L}^d, \mu_2 = f_2 \mathcal{L}^d \in \mathcal{P}_2 \cap \mathcal{P}^{ac}$, let T be a map such that $[\text{id} \times T]_{\#} \mu_1$ is an optimal transport of them. For a C^2 -function Ψ on \mathbb{R}^d such that $\text{Hess } \Psi$ is bounded below by some number K , we have*

$$\int_{\mathbb{R}^d} (f_2 - f_1) \Psi \, d\mathcal{L}^d \geq \int_{\mathbb{R}^d} \langle \nabla \Psi, T - \text{id} \rangle f_1 \, d\mathcal{L}^d + \frac{K}{2} W_2(\mu_1, \mu_2)^2. \tag{5.1}$$

If we moreover assume that f_1 has a weak derivative, then we have

$$\begin{aligned} \int_{\mathbb{R}^d} [f_2 \ln_q(f_2) - f_1 \ln_q(f_1)] \, d\mathcal{L}^d &\geq \int_{\mathbb{R}^d} \langle T - \text{id}, \nabla (f_1^{2-q}) \rangle \, d\mathcal{L}^d \\ &= \int_{\mathbb{R}^d} (2 - q) \langle T - \text{id}, \nabla \ln_q(f_1) \rangle \, d\mu_1. \end{aligned} \tag{5.2}$$

Proof. Since the assumption $\text{Hess } \Psi \geq K$ provides

$$\Psi(T(x)) - \Psi(x) \geq \langle \nabla \Psi(x), T(x) - x \rangle + \frac{K}{2} |x - T(x)|^2,$$

we obtain (5.1) by integrating it with respect to μ_1 , where we use the optimality of $[\text{id} \times T]_{\#} \mu_1$.

We next prove (5.2). Since the equality is trivial, we only prove the inequality. Due to Theorem 2.1 and (2.1), there exists a convex function φ such that

$$\nabla\varphi = T, \quad \text{Hess}\varphi = dT, \quad f_1 = f_2(T) \det(dT)$$

hold for μ_1 -almost everywhere. Let \mathcal{U} be the maximal subset where φ has the second derivative. Then the change of variable formula for $y = T(x)$ implies

$$\int_{\mathbb{R}^d} f_2 \ln_q(f_2) d\mathcal{L}^d = \int_{\mathcal{U}} f_1 \ln_q(f_2(T)) d\mathcal{L}^d$$

and

$$\int_{\mathbb{R}^d} [f_2 \ln_q(f_2) - f_1 \ln_q(f_1)] d\mathcal{L}^d = \int_{\mathcal{U}} [\ln_q(f_2(T)) - \ln_q(f_1)] f_1 d\mathcal{L}^d.$$

Fix $x \in \mathcal{U}$ and set

$$b(t) := \ln_q\left(\frac{f_1(x)}{\det[I_d + t(\text{Hess}\varphi(x) - I_d)]}\right),$$

which is convex for $t \in [0, 1]$ (see [11, Theorem 2.2]). Hence we have

$$b(1) - b(0) = \ln_q(f_2(T(x))) - \ln_q(f_1(x)) \geq b'(0) = -f_1^{1-q}(x) \Delta\left(\varphi(x) - \frac{|x|^2}{2}\right)$$

and by integrating it with respect to $\mu_1 = f_1 \mathcal{L}^d$, we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} [f_2 \ln_q(f_2) - f_1 \ln_q(f_1)] d\mathcal{L}^d &\geq - \int_{\mathcal{U}} f_1^{1-q}(x) \Delta\left(\varphi(x) - \frac{|x|^2}{2}\right) f_1(x) d\mathcal{L}^d(x) \\ &\geq - \int_{\mathbb{R}^d} f_1^{1-q}(x) \Delta_{\mathcal{D}'}\left(\varphi(x) - \frac{|x|^2}{2}\right) f_1(x) d\mathcal{L}^d(x) \\ &= \int_{\mathbb{R}^d} \langle \nabla(f_1)^{2-q}, \nabla\varphi - \text{id} \rangle d\mathcal{L}^d, \end{aligned}$$

where $\Delta_{\mathcal{D}'}$ is the distributional Laplacian. In the second inequality, we use the Aleksandrov theorem [10] which states that $\Delta_{\mathcal{D}'}\varphi$ coincides with $\Delta\varphi$ on \mathcal{U} , and the non-negativity of $\Delta_{\mathcal{D}'}\varphi$ on the interior of the domain of φ , which is derived from the convexity of φ . Thus the lemma is proved. \square

We now give a criterion for a probability measure to satisfy the q -logarithmic Sobolev inequality and the q -Talagrand inequality. For $\nu \in \mathcal{P}_2$, we set

$$\mathcal{P}_2^{\text{ac}}(\nu) := \{\mu \in \mathcal{P}_2 \cap \mathcal{P}^{\text{ac}} \mid \mu \text{ is absolutely continuous with respect to } \nu\}.$$

Proposition 5.2. For $\nu := \exp_q(-\Psi)\mathcal{L}^d \in \mathcal{P}_2$ with $q \in Q_d$, we assume that Ψ is C^2 and that $\text{Hess}\Psi$ is bounded below by a positive number λ .

(1) The q -logarithmic Sobolev inequality

$$H_q(\mu|\nu) \leq \frac{1}{2\lambda} I_q(\mu|\nu) \tag{5.3}$$

holds for all $\mu \in \mathcal{P}_2^{\text{ac}}(\nu)$ such that the density of μ has a first weak derivative.

(2) The q -Talagrand inequality

$$W_2(\mu, \nu) \leq \sqrt{\frac{2}{\lambda} H_q(\mu|\nu)} \tag{5.4}$$

holds for all $\mu \in \mathcal{P}_2^{\text{ac}}(\nu)$.

Proof. (1) Lemma 5.1 with $(f_1, f_2) = (d\mu/d\mathcal{L}^d, \exp_q(-\Psi))$ provides

$$\begin{aligned} -H_q(\mu|v) &= \frac{-1}{2-q} \int_{\mathbb{R}^d} [f_1 \ln_q(f_1) - f_2 \ln_q(f_2) - (2-q)\ln_q(f_2)(f_1 - f_2)] d\mathcal{L}^d \\ &= \frac{-1}{2-q} \int_{\mathbb{R}^d} [f_1 \ln_q(f_1) - f_2 \ln_q(f_2) + (2-q)\Psi(f_1 - f_2)] d\mathcal{L}^d \\ &\geq \int_{\mathbb{R}^d} \left[(T - \text{id}, \nabla(\ln_q(f_1) + \Psi)) + \frac{\lambda}{2} |\text{id} - T|^2 \right] f_1 d\mathcal{L}^d \\ &\geq - \int_{\mathbb{R}^d} \frac{1}{2\lambda} |\nabla(\ln_q(f_1) - \ln_q(f_2))|^2 f_1 d\mathcal{L}^d \\ &= -\frac{1}{2\lambda} I_q(\mu|v), \end{aligned}$$

where we use the assumption $\mu \in \mathcal{P}_2^{\text{ac}}(v)$ in the second line and complete the square in the second inequality.

(2) Setting $(\mu_1, \mu_2) = (v, \mu)$ and adding up the inequality (5.1), (5.2), we have

$$\frac{\lambda}{2} W_2(\mu, v)^2 \leq H_q(\mu|v). \quad \square$$

Let us demonstrate that any q -Gaussian measure provides a nontrivial pair which attains equality in (5.3) and (5.4).

Corollary C. For $V \in \text{Sym}^+(d, \mathbb{R})$, let σ be the largest eigenvalue of V and v^* be the corresponding eigenvector. Then for any $q > 1$ and $v \in \mathbb{R}^d$, a pair $(\mu, \nu) = (N_q(v + v^*, V), N_q(v, V))$ attains equality in the inequalities (5.3) and (5.4) for λ satisfying

$$\lambda\sigma = C_0^{1-q} C_1(\det V)^{-\frac{1-q}{2}}.$$

Remark 5.3. For the function Ψ satisfying $\exp_q(-\Psi) = n_q(v, V)$, we have $\text{Hess } \Psi = C_0^{1-q} C_1(\det V)^{-(1-q)/2} V^{-1}$, that is, the Hessian of Ψ is bounded below by λ . The assumption $q > 1$ guarantees $\mathcal{P}_2^{\text{ac}}(v) = \mathcal{P}_2 \cap \mathcal{P}^{\text{ac}}$.

Proof of Corollary C. By the proof of Proposition 5.2, equality holds for a pair (μ, ν) in (5.3) if and only if the inequalities in (5.1), (5.2) are equalities and we have

$$\lambda(x - T(x)) = \nabla \left(\ln_q \left(\frac{d\mu}{d\mathcal{L}^d} \right) + \Psi(x) \right) \tag{5.5}$$

for μ -almost everywhere, where T is an optimal transport map of μ and ν . Similarly, equality holds for a pair (μ, ν) in (5.4) if and only if the inequalities in (5.1) and (5.2) are equalities. To have equality in (5.2), T must be a translation map, that is, $T(x) = x + u$ for some $u \in \mathbb{R}^d$ (see [11, Theorem 2.2]). In the case of $T(x) = x + u$, equality in each of (5.1) and (5.5) is equivalent to that $\nabla \Psi(T(x)) - \nabla \Psi(x) = \lambda u$ holds for μ -almost everywhere. The pair $(N_q(v + tv^*, V), N_q(v, V))$ satisfies these conditions, hence we have equality in (5.3) and (5.4). \square

Remark 5.4. The proof of Corollary C is needed in order to consider conditions to establish equality in (5.3) and (5.4). We directly show that the pair $(\mu, \nu) = (N_q(v + v^*, V), N_q(v, V))$ given in Corollary C attains equality in (5.3) and (5.4) by computing

$$H_q(\mu|v) = \frac{\lambda}{2} |v^*|^2, \quad I_q(\mu|v) = \lambda^2 |v^*|^2, \quad W_2(\mu, \nu) = |v^*|.$$

In this case, the both sides of (5.3) coincide with $\lambda|v^*|/2$ and the both sides of (5.4) are equal to $|v^*|$.

Remark 5.5. Since the square root of H_q behaves like the distance function in information geometry, the q -Talagrand inequality provides the comparison between the two different geometries, namely, Wasserstein geometry and information geometry.

We next provide a criterion for a probability measure to satisfy the q -HWI inequality using not Lemma 5.1 but the geodesic equation. For a geodesic $\{\mu_t\}_{t \in [0,1]}$ in the Wasserstein space, there exists a family $\{\Phi_t\}_{t \in [0,1]}$ of functions such that

$$\begin{cases} \frac{\partial}{\partial t} \mu_t + \operatorname{div}(\mu_t \nabla \Phi_t) = 0, \\ \frac{\partial}{\partial t} \Phi_t + \frac{1}{2} |\nabla \Phi_t|^2 = 0, \end{cases} \tag{5.6}$$

where the function Φ_t serves as a “tangent vector field” on the Wasserstein space (a heuristics can be found in [15] and also see [23,24]). In short, the relation between this “tangent vector field” Φ_t and an optimal transport map T from μ_0 to μ_1 is

$$\nabla \Phi_t = [T - \operatorname{id}] \circ (T_t^{-1}), \quad T_t := \operatorname{id} + t[T - \operatorname{id}].$$

Proposition 5.6. For $\nu := \exp_q(-\Psi) \mathcal{L}^d \in \mathcal{P}_2$ with $q \in Q_d$ and $q \leq (d + 1)/d$, we assume that Ψ is C^2 and that $\operatorname{Hess} \Psi$ is bounded below by some number K . Suppose that the support of ν is convex. Then the q -HWI inequality

$$H_q(\mu|\nu) \leq W_2(\mu, \nu) \sqrt{I_q(\mu|\nu)} - \frac{K}{2} W_2(\mu, \nu)^2 \tag{5.7}$$

holds for any $\mu \in \mathcal{P}_2^{\operatorname{ac}}(\nu)$ such that the density of μ has a first weak derive.

Remark 5.7. If K is positive, then the support of ν is automatically convex (see [13, Lemma 2.5]). The convexity of the support of ν provides the convexity of $\mathcal{P}_2^{\operatorname{ac}}(\nu)$ in the Wasserstein space (see [8, Corollary 4.3]).

Proof of Proposition 5.6. Let $\{\mu_t = f_t \mathcal{L}^d\}_{t \in [0,1]}$ be a geodesic from $\mu \in \mathcal{P}_2^{\operatorname{ac}}(\nu)$ to ν and $\{\Phi_t\}_{t \in [0,1]}$ be the function satisfying (5.6). We calculate

$$\frac{d}{dt} H_q(\mu_t|\nu) = \int_{\mathbb{R}^d} \langle \nabla \Phi_t, \nabla [\ln_q(f_t) - \ln_q(f_1)] \rangle d\mu_t, \tag{5.8}$$

$$\frac{d^2}{dt^2} H_q(\mu_t|\nu) = \int_{\mathbb{R}^d} \left[\operatorname{Hess} \Psi(\nabla \Phi_t, \nabla \Phi_t) + \frac{f_t^{1-q}}{2-q} \left((1-q)(\Delta \Phi_t)^2 + \sum_{ij} (\operatorname{Hess} \Phi_t)_{ij}^2 \right) \right] d\mu_t, \tag{5.9}$$

$$\frac{d}{dt} \int_{\mathbb{R}^d} |\nabla \Phi_t|^2 d\mu_t = 0, \tag{5.10}$$

where we use the condition that $-\ln_q(\exp_q(-\Psi)) = \Psi$ holds on the support of $\mu_t \in \mathcal{P}_2^{\operatorname{ac}}(\nu)$. Combining the inequality

$$(1-q)(\Delta \Phi_t)^2 + \sum_{ij} (\operatorname{Hess} \Phi_t)_{ij}^2 \geq \left(1-q + \frac{1}{d}\right) (\Delta \Phi_t)^2 \geq 0 \tag{5.11}$$

provided by the Cauchy–Schwarz inequality and the assumption $q \leq (d + 1)/d$, with the assumption $\operatorname{Hess} \Psi \geq K$, we deduce from (5.9) and (5.10) that

$$\frac{d^2}{dt^2} H_q(\mu_t|\nu) \geq \int_{\mathbb{R}^d} K |\nabla \Phi_t|^2 d\mu_t = K W_2(\mu_0, \mu_1)^2 = K W_2(\mu, \nu)^2. \tag{5.12}$$

It follows that

$$\begin{aligned}
 -H_q(\mu|v) &= H_q(\mu_1|v) - H_q(\mu_0|v) \\
 &= \int_0^1 \left[\int_0^t \frac{d^2}{ds^2} H_q(\mu_s|v) ds + \frac{d}{ds} \Big|_{s=0} H_q(\mu_s|v) \right] dt \\
 &\geq \int_0^1 \left[\int_0^t K W_2(\mu_0, \mu_1)^2 ds + \frac{d}{ds} \Big|_{s=0} H_q(\mu_s|v) \right] dt \\
 &= \frac{K}{2} W_2(\mu_0, \mu_1)^2 + \int_{\mathbb{R}^d} \langle \nabla \Phi_0, \nabla [\ln_q(f_0) - \ln_q(f_1)] \rangle d\mu_t \\
 &\geq \frac{K}{2} W_2(\mu_0, \mu_1)^2 - \sqrt{\int_{\mathbb{R}^d} |\nabla \Phi_0|^2 d\mu_t \int_{\mathbb{R}^d} |\nabla [\ln_q(f_0) - \ln_q(f_1)]|^2 d\mu_t} \\
 &= \frac{K}{2} W_2(\mu, v)^2 - W_2(\mu, v) \sqrt{I_q(\mu|v)},
 \end{aligned} \tag{5.13}$$

where the third line follows from (5.12) and the forth line follows from (5.8). In the fifth line, we apply the Cauchy–Schwarz inequality. This concludes the proof of Proposition 5.6. \square

The pair of q -Gaussian measure given in Corollary C also attains equality in the inequality (5.7).

Corollary D. For $V \in \text{Sym}^+(d, \mathbb{R})$, let σ be the largest eigenvalue of V and v^* be the corresponding eigenvector. Then for any $q > 1$ and $v \in \mathbb{R}^d$, a pair $(\mu, \nu) = (N_q(v + v^*, V), N_q(v, V))$ attains equality in the inequality (5.7) for λ satisfying

$$\lambda \sigma = C_0^{1-q} C_1 (\det V)^{-\frac{1-q}{2}}.$$

Proof. To establish equality in (5.7), $\text{Hess } \Psi(\nabla \Phi_t, \nabla \Phi_t) = K |\nabla \Phi_t|^2$ holds and the inequalities in (5.11), (5.13) need to be equalities, which is equivalent to $\nabla \Phi_t \equiv u$, where u is the difference between the mean of ν and μ . The given pair satisfies these conditions and we have equality in (5.7). \square

We finally give a new relation between the q -logarithmic inequality and the q -Talagrand inequality. To do this, we deal with the following evolution equation given by

$$\begin{cases} \frac{\partial}{\partial t} \rho = \frac{1}{2-q} \Delta(\rho^{2-q}) + \text{div}(\rho \nabla \Psi), & x \in \Omega, \\ \rho = 0, & x \in \Omega^c, \end{cases} \tag{5.14}$$

where Ω is an open set defined by

$$\Omega := \{x \in \mathbb{R}^d \mid (1 - q)\Psi(x) < 1\}.$$

For technical reasons, we need to require the solutions ρ of (5.14) to satisfy the following conditions (I), (II) and (III).

- (I) ρ is non-negative and smooth.
- (II) ρ conserves the total mass and has a finite third moment, that is,

$$\int_{\mathbb{R}^d} \rho d\mathcal{L}^d = \int_{\Omega} \rho d\mathcal{L}^d = 1, \quad \int_{\mathbb{R}^d} |x|^3 \rho d\mathcal{L}^d < \infty.$$

(III) For $\xi_t(x) := \ln_q(\rho) - \ln_q(\exp_q(-\Psi))$, there exists a locally bounded function $a(t)$ such that

- (a) $|\nabla(\xi_t(x) - \xi_t(y))| \leq a(t)|x - y|$.
- (b) $|\nabla(\xi_t^2(x))| \leq a(t)(1 + |x|^3)$.

For a function Ψ , we set

$$\mathcal{P}(\Psi) := \{f \mathcal{L}^d \in \mathcal{P}^{\text{ac}} \mid \text{the solution to (5.14) with initial data } f \text{ satisfies (I)–(III)}\}.$$

Remark 5.8. Let $\Psi(x) = \lambda|x|^2/2$ with some $\lambda > 0$. If $1 < q < (d + 5)/(d + 3)$, then any q -Gaussian measure belongs to $\mathcal{P}(\Psi)$ (any q -Gaussian measure has a finite third moment if $q < (d + 5)/(d + 3)$). For $q < 1$, $N_q(0, aI_d) \in \mathcal{P}(\Psi)$ if $\lambda^{2\alpha} a < (C_0^{1-q} C_1)^{2\alpha}$, which ensures that the support of $N_q(0, aI_d)$ is contained in Ω .

Theorem E. For $\nu := \exp_q(-\Psi) \mathcal{L}^d \in \mathcal{P}_2$ with $q \in Q_d$, we assume that Ψ is C^2 and that $\text{Hess } \Psi$ is bounded below by a positive number K . If there exists some positive number λ such that

$$H_q(\mu|\nu) \leq \frac{1}{2\lambda} I_q(\mu|\nu)$$

holds for any $\mu \in \mathcal{P}_2^{\text{ac}}(\nu) \cap \mathcal{P}(\Psi)$, then we also have

$$W_2(\mu, \nu) \leq \sqrt{\frac{2}{\lambda} I_q(\mu|\nu)}$$

for any $\mu \in \mathcal{P}_2^{\text{ac}}(\nu) \cap \mathcal{P}(\Psi)$.

Remark 5.9. Proposition 5.2 says that the condition $\text{Hess } \Psi \geq K > 0$ implies that the q -Talagrand inequality (5.4) holds with the constant K . However, λ is independent of K in Theorem E and the estimate in the q -Talagrand inequality is improved if $\lambda > K$.

The basic strategy for proving Theorem E is similar to the proof of [15, Theorem 1], where the key ingredients are some variations along the flow (5.14) and the convergence of the solution in the sense of H_q and W_2 .

For $\mu = f \mathcal{L}^d$, let f_t be the solution to (5.14) with the initial data $f_0 = f$. Then f_t satisfies

$$\frac{\partial}{\partial t} f_t = \frac{1}{2-q} \Delta(f_t^{2-q}) + \text{div}(f_t \nabla \Psi) = \text{div}(f_t \nabla \xi_t)$$

and the conditions (I) and (II) guarantee $\mu_t := f_t \mathcal{L}^d \in \mathcal{P}_2^{\text{ac}}(\nu)$. We first consider the variations of $H_q(\mu_t|\nu)$ and $W_2(\mu, \mu_t)$ along the solution to (5.14).

Lemma 5.10.

$$\frac{d}{dt} H_q(\mu_t|\nu) = -I_q(\mu_t|\nu), \tag{5.15}$$

$$\frac{d^+}{dt} W_2(\mu, \mu_t) = \limsup_{s \downarrow 0} \frac{W_2(\mu, \mu_{t+s}) - W_2(\mu, \mu_t)}{s} \leq \sqrt{I_q(\mu_t|\nu)}. \tag{5.16}$$

Proof. Set $g := \exp_q(-\Psi)$. For any smooth function η with compact support, we have

$$\begin{aligned} & \frac{d}{dt} \left(\frac{1}{2-q} \int_{\mathbb{R}^d} [f_t \ln_q(f_t) - g \ln_q(g) - (2-q) \ln_q(g)(f_t - g)] \eta d\mathcal{L}^d \right) \\ &= - \int_{\mathbb{R}^d} f_t \left\langle \nabla \xi_t, \nabla \left[\left(\xi_t + \frac{1}{2-q} \right) \eta \right] \right\rangle d\mathcal{L}^d \\ &= - \frac{1}{2} \int_{\mathbb{R}^d} \langle \nabla \eta, \nabla(\xi_t^2) \rangle d\mu_t - \frac{1}{2-q} \int_{\mathbb{R}^d} \langle \nabla \eta, \nabla \xi_t \rangle d\mu_t - \int_{\mathbb{R}^d} \eta |\nabla \xi_t|^2 d\mu_t, \end{aligned}$$

where we use integration by parts. We choose a sequence $\{\eta_n\}$ of smooth functions with compact support satisfying

- η_n is uniformly bounded and converges pointwise to 1,
- $\nabla \eta_n$ is uniformly bounded and converges pointwise to 0.

The Dominated Convergence Theorem with the condition (III) yields

$$\begin{aligned} & \frac{d}{dt} \left(\frac{1}{2-q} \int_{\mathbb{R}^d} [f_t \ln_q(f_t) - g \ln_q(g) - (2-q) \ln_q(g)(f_t - g)] d\mathcal{L}^d \right) \\ &= - \int_{\mathbb{R}^d} |\nabla [f_t \ln_q(f_t) - g \ln_q(g)]|^2 d\mu_t, \end{aligned}$$

proving (5.15).

We next compute the variation of $W_2(\mu, \mu_t)$. Due to the triangle inequality for the Wasserstein distance function

$$|W_2(\mu, \mu_{t+s}) - W_2(\mu, \mu_t)| \leq W_2(\mu_t, \mu_{t+s}),$$

we only need to show

$$\limsup_{s \downarrow 0} \frac{1}{s} W_2(\mu_t, \mu_{t+s}) \leq \sqrt{I_q(\mu_t | \nu)}.$$

The condition (III)(a) guarantees the existence of a family $\{\varphi_s\}_{s \in [0, \varepsilon]}$ of diffeomorphisms on Ω such that

$$\frac{\partial \varphi_s(x)}{\partial s} = -(\nabla \xi_{t+s}) \circ \varphi_s(x) \quad \text{and} \quad \varphi_0 = \text{id}$$

for small enough $\varepsilon > 0$. First we prove that $[\varphi_s]_{\#} \mu_t = \mu_{t+s}$, that is,

$$\int_{\mathbb{R}^d} \eta(\varphi_s^{-1}) d\mu_{t+s} = \int_{\mathbb{R}^d} \eta d\mu_t \tag{5.17}$$

for all smooth functions η with compact support. For $\eta_s := \eta \circ \varphi_s^{-1}$, we have

$$\begin{aligned} 0 &= \frac{\partial \eta}{\partial s} = \frac{\partial(\eta_s \circ \varphi_s)}{\partial s} = \frac{\partial \eta_s}{\partial s}(\varphi_s) + \left\langle \nabla \eta_s \circ \varphi_s, \frac{\partial \varphi_s}{\partial s} \right\rangle \\ &= \left(\frac{\partial \eta_s}{\partial s} - \langle \nabla \eta_s, \nabla \xi_{t+s} \rangle \right) \circ \varphi_s, \end{aligned}$$

which implies

$$\frac{d}{ds} \int_{\mathbb{R}^d} \eta_s d\mu_{t+s} = \int_{\mathbb{R}^d} \left[\frac{\partial \eta_s}{\partial s} - \langle \nabla \eta_s, \nabla \xi_{t+s} \rangle \right] d\mu_{t+s} = 0,$$

where we use the assumption f_t which is the solution to (5.14). It follows that

$$\int_{\mathbb{R}^d} \eta(\varphi_s^{-1}) d\mu_{t+s} = \int_{\mathbb{R}^d} \eta_s d\mu_{t+s} = \int_{\mathbb{R}^d} \eta(\varphi_0^{-1}) d\mu_t = \int_{\mathbb{R}^d} \eta(y) d\mu_t,$$

which is (5.17). Hence $[\text{id} \times \varphi_s]_{\#} \mu_t$ is a transport plan of μ_t and μ_{t+s} and by definition of Wasserstein distance function, we have

$$\frac{1}{s} W_2(\mu_t, \mu_{t+s}) \leq \sqrt{\int_{\mathbb{R}^d} \frac{|x - \varphi_s(x)|^2}{s^2} d\mu_t}.$$

The condition (III)(a) ensures that $(|x - \varphi_s(x)|/s)^2$ is uniformly bounded above by an integrable function with respect to μ_t for $s \in (0, \varepsilon]$, and converges to the integrable function $|\nabla \xi_t|^2$ as $s \rightarrow 0$. The Dominated Convergence Theorem yields

$$\limsup_{s \downarrow 0} \frac{1}{s} W_2(\mu_t, \mu_{t+s}) \leq \sqrt{\int_{\mathbb{R}^d} |\nabla \xi_t|^2 d\mu_t} = \sqrt{I_q(\mu_t|v)},$$

which is the desired result. \square

We next investigate the asymptotic behavior of μ_t in terms of H_q and W_2 .

Lemma 5.11.

$$H_q(\mu_t|v) \xrightarrow{t \uparrow \infty} 0, \quad W_2(\mu, \mu_t) \xrightarrow{t \uparrow \infty} W_2(\mu, v).$$

Proof. From (5.15) and the assumption that the q -logarithmic Sobolev inequality (5.3) holds for $\mu \in \mathcal{P}_2^{\text{ac}}(v) \cap \mathcal{P}(\Psi)$, we deduce

$$\frac{d}{dt} H_q(\mu_t|v) = -I_q(\mu_t|v) \leq -2\lambda H_q(\mu_t|v).$$

Integrating both sides of the inequality above, we have

$$H_q(\mu_t|v) \leq e^{-2\lambda t} H_q(\mu|v) \xrightarrow{t \uparrow \infty} 0.$$

Proposition 5.2(2) with the assumption $\text{Hess } \Psi \geq K$ provides

$$W_2(\mu_t, v) \leq \sqrt{\frac{2}{K} H_q(\mu_t|v)} \xrightarrow{t \uparrow \infty} 0.$$

Combining this with the triangle inequality, we obtain the conclusion. \square

Proof of Theorem E. For the function

$$\psi(t) := W_2(\mu, \mu_t) + \sqrt{\frac{2}{\lambda} H_q(\mu_t|v)},$$

we have

$$\frac{d^+}{dt} \psi(t) \leq \sqrt{I_q(\mu_t|v)} - \frac{I_q(\mu_t|v)}{\sqrt{2\lambda H_q(\mu_t|v)}} \leq 0,$$

where we apply Lemma 5.10 in the first inequality and the second inequality follows from (5.3). Combining the monotonicity of ψ with Lemma 5.11, we obtain the desired inequality, that is,

$$W_2(\mu, v) = \lim_{t \uparrow \infty} \psi(t) \leq \psi(0) = \sqrt{\frac{2}{\lambda} H_q(\mu|v)}. \quad \square$$

Acknowledgements

This paper contains a part of the author’s doctoral dissertation at Tohoku University, and she wishes here to thank her supervisor Sumio Yamada, for his numerous advice, support and encouragement. She would also like to express her gratitude to Takashi Shioya and Shin-ichi Ohta for many valuable suggestions and discussions. She also thanks Hiroshi Matsuzoe and Atsumi Ohara for their explanations of information geometry. She acknowledges Kazuhiro Ishige and Tatsuki Kawakami for their comments. She thanks the anonymous referees for valuable remarks which lead to several crucial revisions.

References

- [1] M. Agueh, N. Ghoussoub, X. Kang, Geometric inequalities via a general comparison principle for interacting gases, *Geom. Funct. Anal.* 14 (1) (2004) 215–244.
- [2] S. Amari, *Differential–Geometrical Methods in Statistics*, Lecture Notes in Statist., vol. 28, Springer-Verlag, New York, 1985.
- [3] S. Amari, H. Nagaoka, *Methods of Information Geometry*, Transl. Math. Monogr., vol. 191, Amer. Math. Soc., Providence, RI, 2000; translated from the 1993 Japanese original by Daishi Harada.
- [4] G.I. Barenblatt, On self-similar motions of a compressible fluid in a porous medium, *Akad. Nauk SSSR Prikl. Mat. Meh.* 16 (1952) 679–698.
- [5] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* 44 (4) (1991) 375–417.
- [6] J.A. Carrillo, A. Jüngel, P.A. Markowich, G. Toscani, A. Unterreiter, Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities, *Monatsh. Math.* 133 (1) (2001) 1–82.
- [7] J.A. Carrillo, R.J. McCann, C. Villani, Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates, *Rev. Mat. Iberoam.* 19 (3) (2003) 971–1018.
- [8] J.A. Carrillo, R.J. McCann, C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media, *Arch. Ration. Mech. Anal.* 179 (2) (2006) 217–263.
- [9] D. Cordero-Erausquin, W. Gangbo, C. Houdré, Inequalities for generalized entropy and optimal transportation, in: *Recent Advances in the Theory and Applications of Mass Transport*, in: *Contemp. Math.*, vol. 353, Amer. Math. Soc., Providence, RI, 2004, pp. 73–94.
- [10] L.C. Evans, R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Stud. Adv. Math., CRC Press, Boca Raton, FL, 1992.
- [11] R.J. McCann, A convexity principle for interacting gases, *Adv. Math.* 128 (1) (1997) 153–179.
- [12] A. Ohara, T. Wada, Information geometry of q -Gaussian densities and behaviors of solutions to related diffusion equations, *J. Phys. A* 43 (2010) 035002, 18 pp.
- [13] S. Ohta, A. Takatsu, Displacement convexity of generalized relative entropies, *Adv. Math.* 228 (2011) 1742–1787.
- [14] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Comm. Partial Differential Equations* 26 (1–2) (2001) 101–174.
- [15] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* 173 (2) (2000) 361–400.
- [16] R.E. Pattle, Diffusion from an instantaneous point source with a concentration-dependent coefficient, *Quart. J. Mech. Appl. Math.* 12 (1959) 407–409.
- [17] A. Takatsu, Wasserstein geometry of Gaussian measures, *Osaka J. Math.* 48 (4) (2011), in press.
- [18] A. Takatsu, Wasserstein geometry of non-linear Fokker–Planck type equations, *Sūrikaiseikikenkyūsho Kōkyūroku* 1671 (2009) 20–36 (in Japanese).
- [19] A. Takatsu, Behaviors of φ -exponential distributions in Wasserstein geometry and an evolution equation, preprint, available at arXiv:1109.6776, 2011.
- [20] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, *J. Stat. Phys.* 52 (1–2) (1988) 479–487.
- [21] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics—Approaching a Complex World*, Springer-Verlag, Berlin, 2009.
- [22] C. Vignat, A. Hero, J. Costa, About closedness by convolution of the Tsallis maximizers, *Phys. A* 340 (1–3) (2004) 147–152.
- [23] C. Villani, *Topics in Optimal Transportation*, Grad. Stud. Math., vol. 58, Amer. Math. Soc., Providence, RI, 2003.
- [24] C. Villani, *Optimal Transport, Old and New*, Grundlehren Math. Wiss., vol. 338, Springer-Verlag, Berlin, 2009.