Advanced Studies in Pure Mathematics ,

pp. 1–20

Optimal transport and Ricci curvature in Finsler geometry

Shin-ichi Ohta

Abstract.

This is a survey article on recent progress (in [Oh3], [OS]) of the theory of weighted Ricci curvature in Finsler geometry. Optimal transport theory plays an impressive role as is developed in the Riemannian case by Lott, Sturm and Villani.

§1. Introduction

A Finsler manifold is a generalization of a Riemannian manifold admitting its tangent spaces being Banach spaces or, more generally, Minkowski spaces. Then a natural question is: Is there a canonical measure associated with each Finsler manifold? Generally speaking, the answer is no. There are several constructive measures, such as the Busemann-Hausdorff measure and the Holmes-Thompson measure, each of which is canonical in some sense (see [AT]). One strategy for dealing with this difficulty is not to choose one specific measure and to consider an arbitrary measure in the first place, like the theory of weighted Riemannian manifolds.

In weighted Riemannian geometry, we decompose given a measure m on a Riemannian manifold M into $m = e^{-\psi} \operatorname{vol}_M$ using the Riemannian volume element vol_M. Then the weight function ψ plays an important role, for instance, the Bakry-Émery tensor $\operatorname{Ric} + \operatorname{Hess} \psi$ is known to behave like the (infinite dimensional) Ricci curvature of this weighted space. In order to follow this line in our Finsler setting, we need a reference measure like vol_M. Our idea is that the reference measure does not live in the manifold, but in the unit sphere bundle. More precisely, we do not fix one reference measure, but choose a reference

2000 *Mathematics Subject Classification.* 53C60, 49Q20, 58J35.

Key words and phrases. Finsler geometry, optimal transport, Ricci curvature, comparison theorem.

Supported in part by the Grant-in-Aid for Young Scientists (B) 20740036.

measure for each given unit tangent vector. At this point we are indebted to Shen's extremely useful interpretation of flag and Ricci curvatures. Then it turns out that many results in (weighted) Riemannian geometry involving Ricci curvature can be translated into Finsler geometry.

One of the highlights stems from the connection with optimal transport theory. The curvature-dimension condition and the N-Ricci curvature bound are equivalent to the lower weighted Ricci curvature bound (Theorem 4.2) as is established in the Riemannian case by Lott, Sturm and Villani. This theorem has fruitful applications in interpolation inequalities, functional inequalities and the concentration of measure phenomenon.

Moreover, beyond just a generalization of the Riemannian case, this result has potential applications in at least two directions. One direction is the approximation of new kinds of spaces. The curvature-dimension condition is stable under the measured Gromov-Hausdorff convergence, but even Banach spaces can not be approximated by Riemannian manifolds with a uniform lower Ricci curvature bound. Therefore the limit of Finsler manifolds covers much wider class of spaces than that of Riemannian manifolds. For instance, Cordero-Erausquin's result on the curvature-dimension condition of Banach spaces (see [Vi2, Theorem in page 908]) is an immediate consequence of Theorem 4.2. The other direction is the connection with Banach space theory. Finsler geometry stands at the intersection of Riemannian geometry and the geometry of Banach spaces, and our results make it possible to study Banach spaces in a differential geometric way.

The article is organized as follows. In Sections 2 and 3, we briefly review the basics of Finsler geometry and optimal transport theory. Then Section 4 is devoted to Lott, Sturm and Villani's curvature-dimension condition (and N-Ricci curvature bound) and its applications. Finally, we discuss heat flow on Finsler manifolds in Section 5.

Throughout the article, without otherwise indicated, (M, F) is a connected, forward complete, n-dimensional C^∞-Finsler manifold with $n \geq 2$, and m is an arbitrary positive C^∞-measure on M.

§2. Finsler geometry

We first review the basics of Finsler geometry as well as recently introduced weighted Ricci curvature. We refer to [BCS], [Sh2], [Oh3] and [OS] for further reading.

2.1. Finsler structures

Let M be a connected, n-dimensional C^∞-manifold and denote by $\pi : TM \longrightarrow M$ the natural projection. Given a local coordinate system $(x^i)_{i=1}^n : U \longrightarrow \mathbb{R}^n$ on an open set $U \subset M$, we will always denote by $(x^i, v^i)_{i=1}^n$ the local coordinate system on $\pi^{-1}(U) \subset TM$ given by, for $v \in \pi^{-1}(U)$,

$$v = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}\Big|_{\pi(v)}.$$

Definition 2.1. (Finsler structures) A C^∞*-Finsler structure* of a C^∞-manifold M is a function $F : TM \longrightarrow [0, \infty)$ satisfying the following conditions:

(1) (Regularity) The function F is C^∞ on $TM \setminus 0$, where 0 stands for the zero section.
(2) (Positive homogeneity of degree 1) For any $v \in TM$ and positive number $\lambda > 0$, we have $F(\lambda v) = \lambda F(v)$.
(3) (Strong convexity) Given a local coordinate system $(x^i)_{i=1}^n$ on $U \subset M$, the $n \times n$ matrix

$$\big(g_{ij}(v)\big) := \left(\frac{1}{2}\frac{\partial^2(F^2)}{\partial v^i \partial v^j}(v)\right) \tag{2.1}$$

is positive-definite at every $v \in \pi^{-1}(U) \setminus 0$.

In other words, each tangent space (T_xM, F) is a *Minkowski space* and F varies smoothly in the horizontal direction. We emphasize that, however, F is not necessarily *absolutely homogeneous* (or *reversible*), namely $F(v) \neq F(-v)$ may happen. It is sometimes helpful to consider the *reverse* of F, $\overleftarrow{F}(v) := F(-v)$, which turns everything around (e.g., distance and geodesics).

For each $v \in T_xM \setminus 0$, the strong convexity gives the inner product

$$g_v\left(\sum_{i=1}^n a_i \frac{\partial}{\partial x^i}, \sum_{j=1}^n b_j \frac{\partial}{\partial x^j}\right) := \sum_{i,j=1}^n a_i b_j g_{ij}(v) \tag{2.2}$$

on T_xM. This g_v approximates F in the sense that the unit sphere of the norm $F_v(w) := \sqrt{g_v(w,w)}$ tangents to that of F at $v/F(v)$ up to the second order. Note that, if F is coming from a Riemannian structure, then g_v always coincides with the original Riemannian structure.

The *Legendre transform* $\mathcal{L}^* : T^*M \longrightarrow TM$ associates each covector $\alpha \in T_x^*M$ with a unique vector $v = \mathcal{L}^*(\alpha) \in T_xM$ such that $F(v) = F^*(\alpha)$ and $\alpha \cdot v = F^*(\alpha)^2$, where F^* stands for the dual

Minkowski norm of F on T_x^*M and $\alpha \cdot v$ is the canonical pairing between T_xM and T_x^*M. We remark that $\mathcal{L}^*|_{T_x^*M}$ is linear only if $F|_{T_xM}$ is an inner product, and $\mathcal{L}^*$ is not differentiable at the origin. For a C^1-function $f : M \longrightarrow \mathbb{R}$, we define the *gradient vector* of f at $x \in M$ as the Legendre transform of its differential, i.e., $\nabla f(x) := \mathcal{L}^*(Df(x)) \in T_xM$. By definition, $\nabla f(x)$ points into the direction in which f increases the most. Observe that

$$f\big(\eta(l)\big) - f\big(\eta(0)\big) \le \int_0^l F\Big(\nabla f\big(\eta(t)\big)\Big)\, dt$$

holds for any unit speed C^1-curve $\eta : [0, l] \longrightarrow M$. Note the difference between $\nabla(-f)$ and $-\nabla f$.

For a C^1-curve $\eta : [0, l] \longrightarrow M$, we define its arclength in a natural way by

$$L(\eta) := \int_0^l F\big(\dot\eta(t)\big)\, dt, \qquad \dot\eta(t) := \frac{d\eta}{dt}(t).$$

Then the corresponding distance function $d : M \times M \longrightarrow [0, \infty)$ is given by $d(x, y) := \inf_\eta L(\eta)$, where the infimum is taken over all C^1-curves η from x to y. Unreversibility of F causes nonsymmetry of d, that is, $d(x, y) \neq d(y, x)$. We also remark that the squared distance function $d(x, \cdot)^2$ is C^2 at x if and only if $F|_{T_xM}$ is an inner product ([Sh1, Proposition 2.2]). Given $x \in M$ and $r > 0$, we define the *forward open ball* of center $x \in M$ and radius $r > 0$ by

$$B^+(x, r) := \{y \in M \mid d(x, y) < r\}.$$

We also define the open ball in T_xM by

$$B^+_{T_xM}(0, r) := \{v \in T_xM \mid F(v) < r\}.$$

We say that a C^∞-curve $\eta : [0, l] \longrightarrow M$ is a *geodesic* if it is locally minimizing and has constant speed. Define the *exponential map* by $\exp_x(v) := \eta(1)$ for $v \in T_xM$ if there is a geodesic $\eta : [0, 1] \longrightarrow M$ with $\dot\eta(0) = v$. We say that (M, F) is *forward complete* if every geodesic $\eta : [0, l] \longrightarrow M$ is extended to a geodesic on $[0, \infty)$, in other words, if $\exp_x$ is defined on entire T_xM. Then it follows from the Hopf-Rinow theorem (cf. [BCS, Theorem 6.6.1]) that every pair of two points in M can be joined by a minimal geodesic.

For each unit vector $v \in T_xM$, let $l(v) \in (0, \infty]$ be the supremum of $l > 0$ such that the geodesic $\exp_x tv$ is minimal for $t \in [0, l]$. If $l(v) < \infty$, then the point $\exp_x(l(v)v)$ is called a *cut point* of x, and the *cut locus* $\mathrm{Cut}(x)$ of x is defined as the set of all cut points of x.

2.2. Flag and Ricci curvatures

Flag curvature is a substitute of sectional curvature in Riemannian geometry. We follow the heuristic introduction due to Shen (see [Sh2, Chapter 6]).

Fix a unit vector $v \in T_xM$ and extend it to a C^∞-vector field on an open neighborhood U of x such that $V(x) = v$ and every integral curve of V is a geodesic. A typical example is $V = \nabla[d(\eta(-\varepsilon), \cdot)]$ for sufficiently small $\varepsilon > 0$, where $\eta : [-\varepsilon, \varepsilon] \longrightarrow M$ is the geodesic with $\dot{\eta}(0) = v$. The vector field V induces the Riemannian structure g_V of U through (2.2). Then the *flag curvature* $\mathcal{K}(v, w)$ of v and a linearly independent vector $w \in T_xM$ is the sectional curvature with respect to g_V of the plane spanned by v and w. (We mean that $\mathcal{K}(v, w)$ is independent of the choice of the vector field V.) Similarly, the (Finsler) *Ricci curvature* $\mathrm{Ric}(v)$ of v is the (Riemannian) Ricci curvature of v with respect to g_V. We remark that the flag curvature $\mathcal{K}(v, w)$ depends not only on the *flag* (the plane spanned by v and w), but also on the *flagpole* v.

In this setting, for instance, the Bonnet-Myers theorem and the Cartan-Hadamard theorem are extended verbatim (cf. [BCS, Theorems 7.7.1, 9.4.1]). We recall several fundamental examples of Finsler manifolds.

Example 2.2. (a) Every Minkowski space $(\mathbb{R}^n, F)$ has flat flag curvature. Here $F : \mathbb{R}^n \longrightarrow [0, \infty)$ is a positively homogeneous, strongly convex function and all tangent spaces are canonically identified with $(\mathbb{R}^n, F)$.

(b) Unit spheres of Minkowski spaces form an interesting family of Finsler manifolds, although the author does not know any good reference. There should be a connection with the geometry of Banach spaces.

(c) The Hilbert metric on a bounded, open convex domain $D \subset \mathbb{R}^n$ with smooth boundary such that $D \cup \partial D$ is strictly convex is known to arise from a Finsler metric of constant negative flag curvature.

(d) The Teichmüller metric on Teichmüller space is arguably the most famous Finsler metric in differential geometry.

(e) If a connected, simply-connected, forward complete Finsler manifold of Berwald type has nonpositive flag curvature, then it is nonpositively curved in the sense of Busemann (due to [KVK, Theorem 7] combined with the Cartan-Hadamard theorem, see also [KK]). Roughly speaking, a Finsler manifold of *Berwald type* is modeled by a single Minkowski space (all tangent spaces are isometric to each other). Such Finsler manifolds form a reasonable class including both Riemannian manifolds and Banach spaces.

2.3. Weighted Ricci curvature

We fix an arbitrary positive C^∞-measure m on M. Given a unit vector $v \in T_xM$, recall that v induces the inner product g_v on T_xM through (2.1) and (2.2). Define the function Ψ on the unit sphere bundle $F^{-1}(1) \subset TM$ by

$$\Psi(v) := \log\left(\frac{\omega_v(B^+_{T_xM}(0,1))}{m_x(B^+_{T_xM}(0,1))}\right),$$

where ω_v and m_x stands for the Lebesgue measures on T_xM induced from g_v and m, respectively. We can rewrite this as $m_x = e^{-\Psi(v)}\omega_v$, so that Ψ is a weight function which lives in $F^{-1}(1)$, not in M. For brevity, we set

$$\partial_v\Psi := \frac{d}{dt}\Big|_{t=0}\Psi\big(\dot\eta(t)\big), \qquad \partial_v^2\Psi := \frac{d^2}{dt^2}\Big|_{t=0}\Psi\big(\dot\eta(t)\big),$$

where $\eta : (-\varepsilon,\varepsilon) \longrightarrow M$ is the geodesic with $\dot\eta(0) = v$. Inspired by the theory of weighted Riemannian manifolds ([BE], [Qi], [Lo1]), we have introduced the following weighted Ricci curvature in [Oh3].

Definition 2.3. (Weighted Ricci curvature) Let (M,F) be an n-dimensional C^∞-Finsler manifold ($n \ge 2$) and m be a positive C^∞-measure on M. Given a unit vector $v \in T_xM$, we define

(i) $\mathrm{Ric}_n(v) := \begin{cases} \mathrm{Ric}(v) + \partial_v^2\Psi & \text{if } \partial_v\Psi = 0, \\ -\infty & \text{otherwise,} \end{cases}$

(ii) $\mathrm{Ric}_N(v) := \mathrm{Ric}(v) + \partial_v^2\Psi - \dfrac{(\partial_v\Psi)^2}{N-n}$ for $N \in (n,\infty)$,

(iii) $\mathrm{Ric}_\infty(v) := \mathrm{Ric}(v) + \partial_v^2\Psi$.

Note that Ric_∞ corresponds to the Bakry-Émery tensor and Ric_N is an analogue of Qian's generalized one. We also remark that bounding Ric_n from below makes sense only when $\partial_v\Psi = 0$ for all v. This is the case of the Busemann-Hausdorff measure on a Finsler manifold of Berwald type. However, the existence of such a measure seems a strong constraint among general Finsler manifolds.

We define some functions for later convenience. For $K \in \mathbb{R}$, $N \in (1,\infty)$ and $r \in (0,\infty)$ ($r \in (0,\pi\sqrt{(N-1)/K})$ if $K > 0$), we define

$$\mathbf{s}_{K,N}(r) := \begin{cases} \sqrt{(N-1)/K}\,\sin(r\sqrt{K/(N-1)}) & \text{if } K > 0, \\ r & \text{if } K = 0, \\ \sqrt{-(N-1)/K}\,\sinh(r\sqrt{-K/(N-1)}) & \text{if } K < 0. \end{cases}$$

In addition, for $t \in (0,1)$, we define

$$(2.3) \qquad \beta^t_{K,N}(r) := \left(\frac{\mathbf{s}_{K,N}(tr)}{t\mathbf{s}_{K,N}(r)} \right)^{N-1}, \qquad \beta^t_{K,\infty}(r) := e^{K(1-t^2)r^2/6}.$$

Arguing as in weighted Riemannian manifolds (using the Riemannian structure g_V induced from $V = \nabla d(x,\cdot)$), we immediately obtain a generalized Bishop-Gromov volume comparison theorem.

Theorem 2.4. ([Oh3]) *Assume that there are constants $K \in \mathbb{R}$ and $N \in [n,\infty)$ such that $\mathrm{Ric}_N(v) \ge K$ holds for all unit vectors $v \in TM$. Then we have $\operatorname{diam} M \le \pi\sqrt{(N-1)/K}$ if $K > 0$ and, for any $x \in M$ and $0 < r < R$ ($\le \pi\sqrt{(N-1)/K}$ if $K > 0$), it holds that*

$$\frac{m(B^+(x,R))}{m(B^+(x,r))} \le \frac{\int_0^R \mathbf{s}_{K,N}(t)^{N-1}\,dt}{\int_0^r \mathbf{s}_{K,N}(t)^{N-1}\,dt}.$$

Theorem 2.4 should be compared with Shen's volume comparison theorem.

Theorem 2.5. ([Sh1, Theorem 1.1], [Sh2, Theorem 16.1.1]) *Assume that there are constants $K \in \mathbb{R}$ and $H \ge 0$ such that we have $\mathrm{Ric}(v) \ge K$ and $\partial_v \Psi \ge -H$ for all unit vectors $v \in TM$. Then we have, for any $x \in M$ and $0 < r < R$ ($\le \pi\sqrt{(n-1)/K}$ if $K > 0$),*

$$\frac{m(B^+(x,R))}{m(B^+(x,r))} \le \frac{\int_0^R e^{tH}\mathbf{s}_{K,n}(t)^{n-1}\,dt}{\int_0^r e^{tH}\mathbf{s}_{K,n}(t)^{n-1}\,dt}.$$

Shen's theorem has a number of topological applications, however, its exponential term e^{tH} is troublesome relative to the estimate in Theorem 2.4.

§3. Optimal transport in Finsler manifolds

In this section, we very briefly recall some basic knowledge of optimal transport theory for later use. We refer to [Vi1] and recent comprehensive [Vi2] for further reading.

3.1. The Monge-Kantorovich problem and the Wasserstein distance

Let (M,F) be a C^∞-Finsler manifold and denote by $\mathcal{P}(M)$ the set of Borel probability measures on M. For simplicity, we will restrict our attention to the subset $\mathcal{P}_c(M) \subset \mathcal{P}(M)$ consisting of compactly supported measures. Given $\mu, \nu \in \mathcal{P}_c(M)$, the Monge problem asks how to find and characterize a map $T : M \longrightarrow M$ attaining the infimum of

$$\int_M d\big(x, T(x)\big)^2 \, d\mu(x) \tag{3.1}$$

among all maps pushing μ forward to ν (we will write it as $T_\sharp \mu = \nu$). The quantity (3.1) can regarded as the cost of transporting μ to ν according to the map T.

Much later on, Kantorovich extended the Monge problem in a symmetric way. Define $\Pi(\mu, \nu) \subset \mathcal{P}(M \times M)$ as the set of *couplings* of (μ, ν), that is to say, $\pi \in \Pi(\mu, \nu)$ if $\pi(A \times M) = \mu(A)$ and $\pi(M \times A) = \nu(A)$ hold for any measurable set $A \subset M$ (in other words, marginals of π are μ and ν). Then the Monge-Kantorovich problem is to find and characterize an *optimal* coupling $\pi \in \Pi(\mu, \nu)$ attaining

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{M \times M} d(x, y)^2 \, d\pi(x, y). \tag{3.2}$$

The square root of the infimum (3.2) is called the (L^2-)*Wasserstein distance* $d_2^W(\mu, \nu)$ from μ to ν. Note that $d_2^W(\mu, \nu)$ is not greater than the infimum of (3.1) in the Monge problem.

It is known that $(\mathcal{P}_c(M), d_2^W)$ is an Alexandrov space of nonnegative curvature if M is a Riemannian manifold of nonnegative sectional curvature. This relation can be generalized to Finsler manifolds by introducing the notion of 2-*uniform smoothness* (see [Oh1], [Oh2], [Sa]).

3.2. Optimal transport/coupling via $(d^2/2)$-convex functions

A function $\varphi : M \longrightarrow \mathbb{R}$ is said to be $(d^2/2)$-*concave* if there is a function $\psi : M \longrightarrow \mathbb{R}$ such that

$$\varphi(x) = \inf_{y \in M} \{d^2(x, y)/2 - \psi(y)\}$$

holds for all $x \in M$. Any $(d^2/2)$-concave function is Lipschitz continuous and twice differentiable a.e. ([Oh2, Theorem 7.4]). We say that φ is $(d^2/2)$-*convex* if $-\varphi$ is $(d^2/2)$-concave. Then the Brenier-McCann characterization of optimal transport states the following. See [Br], [Mc] for original work in Euclidean spaces and Riemannian manifolds.

Theorem 3.1. ([Oh3]) *For any $\mu, \nu \in \mathcal{P}_c(M)$ such that μ is absolutely continuous, there exists a $(d^2/2)$-convex function $\varphi : M \longrightarrow \mathbb{R}$ such that the map $T(x) := \exp_x(\nabla\varphi(x))$ is a unique optimal transport* *from μ to ν in the sense that $\pi := (\mathrm{Id}_M \times T)_\sharp \mu$ is a unique optimal coupling of (μ, ν). Furthermore, the curve $(\mu_t)_{t\in[0,1]}$ given by $\mu_t = (T_t)_\sharp \mu$ with $T_t(x) = \exp_x(t\nabla\varphi(x))$ is a unique minimal geodesic from μ to ν in $(\mathcal{P}_c(M), d_2^W)$.*

Thus the map T_1 is a unique optimal tansport from μ to ν.

§4. The curvature-dimension condition and its applications

This section is devoted to a main topic of the article, the equivalence between the lower weighted Ricci curvature bound and Lott, Sturm and Villani's curvature-dimension condition as well as N-Ricci curvature bound. See [Oh3] for more details.

4.1. The curvature-dimension condition

In order to define the N-Ricci curvature bound, we recall the important classes of *displacement convex* functions (see [LV1], [LV2] for details). For $N \in [1, \infty)$, denote by $\mathcal{DC}_N$ the set of continuous convex functions $U : [0, \infty) \longrightarrow \mathbb{R}$ such that $U(0) = 0$ and that the function $\varphi(s) := s^N U(s^{-N})$ is convex on $(0, \infty)$. We similarly define $\mathcal{DC}_\infty$ as the set of continuous convex functions $U : [0, \infty) \longrightarrow \mathbb{R}$ such that $U(0) = 0$ and that $\varphi(s) := e^s U(e^{-s})$ is convex on $\mathbb{R}$. For an absolutely continuous measure $\mu = \rho m \in \mathcal{P}(M)$, we define $U_m(\mu) := \int_M U(\rho)\, dm$. We remark that $\mathcal{DC}_{N'} \subset \mathcal{DC}_N$ if $N < N'$.

The most important element of $\mathcal{DC}_N$ is $U(r) = Nr(1-r^{-1/N})$ which derives the *Rényi entropy* $U_m(\rho m) = N - N\int_M \rho^{1-1/N}\, dm$. Letting N go to infinity provides $U(r) = r\log r \in \mathcal{DC}_\infty$ as well as the *relative entropy* $\mathrm{Ent}_m(\rho m) := U_m(\rho m) = \int_M \rho \log \rho\, dm$.

By Theorem 3.1, given two absolutely continuous measures $\mu_0, \mu_1 \in \mathcal{P}_c(M)$, there is a unique minimal geodesic $(\mu_t)_{t\in[0,1]}$ from μ_0 to μ_1 in the Wasserstein space $(\mathcal{P}_c(M), d_2^W)$. Moreover, every μ_t is absolutely continuous (see [Oh3]). Recall (2.3) for the definition of the function $\beta^t_{K,N}$.

Definition 4.1. (N-Ricci curvature bounds, [LV1], [LV2]) For $K \in \mathbb{R}$ and $N \in (1, \infty]$, we say that (M, F, m) has *N-Ricci curvature bounded below by K* if, for any two absolutely continuous probability measures $\mu_0 = \rho_0 m, \mu_1 = \rho_1 m \in \mathcal{P}_c(M)$, $U \in \mathcal{DC}_N$ and for any $t \in (0, 1)$, it holds that

$$U_m(\mu_t) \le (1-t)\int_{M\times M} \frac{\beta_{K,N}^{1-t}(d(x,y))}{\rho_0(x)} U\biggl(\frac{\rho_0(x)}{\beta_{K,N}^{1-t}(d(x,y))}\biggr)\,d\pi(x,y) \\ + t\int_{M\times M} \frac{\beta_{K,N}^{t}(d(x,y))}{\rho_1(y)} U\biggl(\frac{\rho_1(y)}{\beta_{K,N}^{t}(d(x,y))}\biggr)\,d\pi(x,y),$$

where $(\mu_t)_{t\in[0,1]}$ is the unique minimal geodesic from μ_0 to μ_1 and π is the unique optimal coupling of (μ_0,μ_1).

In the particular case of $K=0$, the above inequality reduces to the convexity of $U_m(\mu_\cdot)$. We usually require only the existence of a minimal geodesic $(\mu_t)_{t\in[0,1]}$ satisfying the above inequality in order to ensure the stability under the (measured Gromov-Hausdorff) convergence of spaces. Thus we implicitly took advantage of the unique existence of minimal geodesics in the above definition.

Sturm's *curvature-dimension condition* $\mathsf{CD}(K,N)$ uses the same inequality as Definition 4.1, but only for the Rényi and the relative entropies ([St2], [St3]). These cases are indeed essential in the sense that an n-dimensional Riemannian manifold (M,vol_M) equipped with the Riemannian volume element satisfies $\mathsf{CD}(K,N)$ if and only if $\mathrm{Ric}\ge K$ and $n\le N$. The K-convexity of the relative entropy considered in [vRS] amounts to $\mathsf{CD}(K,\infty)$.

Lott, von Renesse, Sturm and Villani's characterization of weighted Ricci curvature bounds ([vRS], [St1], [St2], [St3], [LV1], [LV2]) is successfully generalized to our Finsler setting.

Theorem 4.2. ([Oh3]) *For $K\in\mathbb{R}$ and $N\in[n,\infty]$, the following conditions are equivalent*:

(i) $\mathrm{Ric}_N(v)\ge K$ *holds for any unit vector* $v\in TM$.

(ii) (M,F,m) *satisfies the curvature-dimension condition* $\mathsf{CD}(K,N)$.

(iii) (M,F,m) *has N-Ricci curvature bounded below by K.*

The proof of the above theorem heavily depends on careful analysis of optimal transport maps developed in [CMS]. Roughly speaking, as an optimal transport T_t goes along the gradient vector field of some $(d^2/2)$-convex function (Theorem 3.1) which is twice differentiable a.e., we can differentiate T_t in the space variable. Then we obtain Jacobi fields and it is natural to control their behavior using the Ricci curvature.

As was briefly mentioned, the N-Ricci curvature bound as well as the curvature-dimension condition are preserved under the measured Gromov-Hausdorff convergence ([St2], [St3], [LV1], [LV2]). Therefore the N-Ricci curvature bound is useful in the investigation of limit spaces of

Finsler manifolds with uniform lower (weighted) Ricci curvature bounds. We refer to [CC] for related, celebrated work in the Riemannian case, and remark that even Banach spaces can not be approximated by Riemannian spaces with a uniform lower Ricci curvature bound.

4.2. Interpolation inequalities

The technique used in the implication (i) $\Rightarrow$ (ii) of Theorem 4.2 has been established in Cordero-Erausquin, McCann and Schmuckenschläger's remarkable work [CMS]. They used it to study several interpolation inequalities on Riemannian manifolds. We refer to [Ga], [Le] for basics and further reading in the Euclidean setting.

The classical Brunn-Minkowski inequality in the Euclidean space $\mathbb{R}^n$ asserts the concavity of the function $t \longmapsto \mathrm{vol}_n((1-t)A+tB)^{1/n}$ of $t \in [0,1]$ for any (nonempty) bounded, measurable sets $A, B \subset \mathbb{R}^n$, where vol_n stands for the Lebesgue measure and

$$(1-t)A + tB := \{(1-t)x + ty \,|\, x \in A,\ y \in B\}.$$

The following nonlinear version is introduced in [St3] (see also [vRS] for the infinite dimensional case) and used to establish the implication (ii) $\Rightarrow$ (i) of Theorem 4.2 in the Riemannian case. Given subsets $A, B \in M$ and $t \in (0,1)$, we denote by $Z_t(A,B)$ the set of points $\eta(t)$ such that $\eta : [0,1] \longrightarrow M$ is a minimal geodesic with $\eta(0) \in A$ and $\eta(1) \in B$.

Theorem 4.3. (Brunn-Minkowski inequality, [St3], [Oh3]) *Suppose that there are constants $K \in \mathbb{R}$ and $N \in [n,\infty]$ such that $\mathrm{Ric}_N(v) \ge K$ holds for all unit vectors $v \in TM$.*

(i) *If $N < \infty$, for any nonempty measurable sets $A, B \subset M$ and $t \in (0,1)$, we have*

$$\begin{aligned} m\big(Z_t(A,B)\big)^{1/N} \ge (1-t) \inf_{x \in A, y \in B} \beta_{K,N}^{1-t}\big(d(x,y)\big)^{1/N} \cdot m(A)^{1/N} \\ + t \inf_{x \in A, y \in B} \beta_{K,N}^{t}\big(d(x,y)\big)^{1/N} \cdot m(B)^{1/N}. \end{aligned}$$

(ii) *If $N = \infty$, for any bounded measurable sets $A, B \subset M$ with $m(A), m(B) > 0$ and $t \in (0,1)$, we have*

$$\begin{aligned} \log m\big(Z_t(A,B)\big) &\ge (1-t)\log m(A) + t \log m(B) \\ &\quad + \frac{K}{2}(1-t)t d_2^W(\mu,\nu)^2, \end{aligned}$$

where $\mu = m(A)^{-1} m|_A$ and $\nu = m(B)^{-1} m|_B$.

The classical Brunn-Minkowski inequality in $\mathbb{R}^n$ admits a functional version called the Prékopa-Leindler inequality. It is also extended to the nonlinear setting.

Theorem 4.4. (Prékopa-Leindler inequality, [CMS], [Oh3]) *Suppose that there are constants $K \in \mathbb{R}$ and $N \in [n, \infty]$ such that $\mathrm{Ric}_N(v) \geq K$ holds for all unit vectors $v \in TM$. Take $t \in (0,1)$, three nonnegative measurable functions $f, g, h : M \longrightarrow [0, \infty)$ and nonempty measurable sets $A, B \subset M$ with $\int_A f\,dm = \int_M f\,dm$ and $\int_B g\,dm = \int_M g\,dm$. If*

$$h(z) \geq \left(\frac{f(x)}{\beta_{K,N}^{1-t}(d(x,y))} \right)^{1-t} \left(\frac{g(y)}{\beta_{K,N}^{t}(d(x,y))} \right)^{t}$$

holds for all $x \in A$, $y \in B$ and $z \in Z_t(x,y)$, then we have $\int_M h\,dm \geq (\int_M f\,dm)^{1-t}(\int_M g\,dm)^t$.

The above Prékopa-Leindler inequality (with $N < \infty$) is the special case ($p = 0$) of the *p-mean inequality* studied in [CMS]. The following (Borell-)Brascamp-Lieb inequality is also the special as well as the strongest case ($p = -1/N$) of the p-mean inequality. See [CMS], [Ga] for more details.

Theorem 4.5. (Brascamp-Lieb inequality, [CMS], [Oh3]) *Assume that there are constants $K \in \mathbb{R}$ and $N \in [n, \infty)$ such that $\mathrm{Ric}_N(v) \geq K$ holds for all unit vectors $v \in TM$. Take three nonnegative measurable functions $f, g, h : M \longrightarrow [0, \infty)$ and measurable sets $A, B \subset M$ with $\int_A f\,dm = \int_B g\,dm = 1$. If there is $t \in (0,1)$ such that*

$$\frac{1}{h(z)^{1/N}} \leq (1-t) \left(\frac{\beta_{K,N}^{1-t}(d(x,y))}{f(x)} \right)^{1/N} + t \left(\frac{\beta_{K,N}^{t}(d(x,y))}{g(y)} \right)^{1/N}$$

holds for all $x \in A$, $y \in B$ and $z \in Z_t(x,y)$, then we have $\int_M h\,dm \geq 1$.

4.3. Functional inequalities and concentration of measures

Connection between the condition like N-Ricci curvature bound and several functional inequalities is first indicated in Otto and Villani's influencial paper [OV], and then systematically investigated by Lott and Villani [LV1], [LV2]. We first recall results in the infinite dimensional case.

Theorem 4.6. ([LV1], [Oh3]) *Let (M, F, m) be a compact Finsler manifold satisfying $m(M) = 1$ and $\mathrm{Ric}_\infty(v) \geq K$ for some $K \in \mathbb{R}$ and all unit vectors $v \in TM$. Then the following hold.*

(i) (Talagrand inequality/Transport cost inequality) *For any absolutely continuous measure $\mu \in \mathcal{P}(M)$, we have*

$$\mathrm{Ent}_m(\mu) \geq \frac{K}{2} \max\left\{d_2^W(m,\mu)^2, d_2^W(\mu,m)^2\right\}.$$

(ii) (Logarithmic Sobolev inequality) *For any Lipschitz continuous function $f : M \longrightarrow \mathbb{R}$ with $\int_M f^2\,dm = 1$, we have*

$$\mathrm{Ent}_m(\mu) \leq 2\biggl(\int_M F(\nabla|f|)^2\,dm\biggr)^{1/2} d_2^W(m,\mu) - \frac{K}{2} d_2^W(m,\mu)^2,$$

where we set $\mu = f^2 m$. In particular, if $K > 0$, then it holds that

$$\mathrm{Ent}_m(\mu) \leq \frac{2}{K} \int_M F(\nabla f)^2\,dm.$$

(iii) (Global Poincaré inequality) *If $K > 0$, then we have, for any Lipschitz continuous function $f : M \longrightarrow \mathbb{R}$ with $\int_M f\,dm = 0$,*

$$\int_M f^2\,dm \leq \frac{1}{K} \int_M F(\nabla f)^2\,dm.$$

The above global Poincaré inequality can be sharpended in the finite dimensional case as follows.

Theorem 4.7. (Lichnerowicz inequality, [LV2], [Oh3]) *Suppose that there are constants $K > 0$ and $N \in [n,\infty)$ such that $\mathrm{Ric}_N(v) \geq K$ holds for all unit vectors $v \in TM$. Then we have, for any Lipschitz continuous function $f : M \longrightarrow \mathbb{R}$ with $\int_M f\,dm = 0$,*

$$\int_M f^2\,dm \leq \frac{N-1}{KN} \int_M F(\nabla f)^2\,dm.$$

Interpolation and functional inequalities are closely related to the concentration of measure phenomenon (see [Le]). For instance, we can generalize Gromov and Milman's Gaussian (normal) concentration result of Riemannian manifolds of positive Ricci curvature ([GM]). For a (symmetric or nonsymmetric) metric space (X,d) equipped with a Borel probability measure μ on X, we define the *concentration function* by

$$\alpha_{(X,d,\mu)}(r) := \sup\left\{1 - \mu\bigl(B^+(A,r)\bigr) \mid A \subset X,\ \mu(A) \geq 1/2\right\}$$

for $r > 0$. Here we set $B^+(A,r) := \{y \in X \mid \inf_{x\in A} d(x,y) < r\}$.

Proposition 4.8. ([Oh3]) *If there is a positive constant $K > 0$ such that $\mathrm{Ric}_\infty(v) \geq K$ holds for all unit vectors $v \in TM$, then we have $\alpha_{(M,F,m)}(r) \leq 2e^{-Kr^2/4}$ for all $r > 0$.*

§5. Laplacian and the heat equation

In this final section, we briefly review the contents of [OS]. For simplicity, we will treat only compact Finsler manifolds.

5.1. Finsler Laplacian

For a vector field V on M differentiable a.e., we define its *divergence* $\operatorname{div} V : M \longrightarrow \mathbb{R}$ through the identity

$$\int_M \phi \operatorname{div} V \, dm = -\int_M V\phi \, dm = -\int_M D\phi \cdot V \, dm$$

for all $\phi \in C^\infty(M)$, where $D\phi(x) \cdot V(x)$ denotes the canonical paring between T_xM and T_x^*M. We remark that $\operatorname{div} V$ depends on the choice of the measure m. If V is differentiable in a local coordinate system $(x^i)_{i=1}^n$, then we have the explicit formula

$$\operatorname{div} V(x) = \sum_{i=1}^n \left\{ \frac{\partial V^i}{\partial x^i}(x) - V^i(x) \frac{\partial \Psi}{\partial x^i}(x) \right\},$$

where $V = \sum_{i=1}^n V^i(\partial/\partial x^i)$ and $m(dx) = e^{-\Psi(x)} dx^1 \cdots dx^n$ in this coordinate. The energy functional $\mathcal{E} : H^1(M) \longrightarrow [0, \infty)$ is defined by

$$\mathcal{E}(u) := \frac{1}{2} \int_M F^*(Du)^2 \, dm = \frac{1}{2} \int_M F(\nabla u)^2 \, dm.$$

For later use, we also set

$$\lambda := \inf \left\{ 2\mathcal{E}(u) \,\middle|\, u \in H^1(M),\ \int_M u \, dm = 0,\ \|u\|_{L^2} = 1 \right\}.$$

It is immediate by virtue of the compactness of M that $\lambda > 0$. Using the inner product g_v in (2.2), we also introduce the constant $\kappa \in (0, 1]$ by

$$\kappa := \inf_{x \in M} \inf_{v, w \in F^{-1}(1) \cap T_xM} \frac{F(w)^2}{g_v(w, w)}.$$

Note that $1/\sqrt{\kappa}$ coincides with the 2-*uniform smoothness constant* studied in [Oh2], and that $\kappa = 1$ holds if and only if (M, F) is Riemannian. We observe by definition that the functional $\mathcal{E}$ is $(\kappa\lambda)$-convex on the set $\{u \in L^2(M; m) \mid \int_M u \, dm = C\}$ for each constant $C \in \mathbb{R}$.

Definition 5.1. (Finsler Laplacian) We define the *Finlser Laplacian* Δ acting on functions $u \in H^1(M)$ formally as $\Delta u := \operatorname{div}(\nabla u)$.

To be precise, Δu is the distributional Laplacian defined through the identity

$$\int_M \phi \Delta u \, dm = -\int_M D\phi \cdot \nabla u \, dm$$

for all $\phi \in C^\infty(M)$.

Recall that, even when u is C^∞, the gradient vector field ∇u is not differentiable at points x with $\nabla u(x) = 0$. Our Finsler Laplacian is a nonlinear operator because the Legendre transform is not linear. Note also that our Laplacian is a negative operator, that is, $\int_M u \Delta u \, dm \le 0$ holds for all $u \in H^1(M)$. The same idea as Theorem 2.4 (using g_V with $V = \nabla d(x, \cdot)$) leads us to the Laplacian comparison theorem.

Theorem 5.2. (Laplacian comparison, [OS]) *Assume that there are constants $K \in \mathbb{R}$ and $N \in [n, \infty)$ such that $\mathrm{Ric}_N(v) \ge K$ holds for all unit vectors $v \in TM$. Then the Laplacian of the distance function $u = d(x, \cdot)$ from any fixed point $x \in M$ satisfies*

$$\Delta u(z) \le (N-1) \frac{\mathbf{s}'_{K,N}(d(x,z))}{\mathbf{s}_{K,N}(d(x,z))}$$

pointwise on $M \setminus (\{x\} \cup \mathrm{Cut}(x))$ and in the sense of distribution on $M \setminus \{x\}$.

5.2. The heat equation

We formulate the heat equation associated with our Finsler Laplacian and show the unique existence of its solution as gradient flow of the energy functional on $L^2(M; m)$.

Definition 5.3. We say that a function u is a (*global*) *solution* to the heat equation $\partial_t u = \Delta u$ on $[0, T] \times M$ if $u \in L^2([0,T], H^1(M)) \cap H^1([0,T], H^{-1}(M))$ and if, for every $t \in [0, T]$ and $\phi \in C^\infty(M)$, it holds that

$$\int_M \phi \partial_t u_t \, dm = -\int_M D\phi \cdot \nabla u_t \, dm,$$

where $u_t(x) := u(t, x)$.

The compactness of M ensures that every solution u to the heat equation is mass preserving, i.e., $\int_M u_t \, dm = \int_M u_0 \, dm$ holds for all $t > 0$.

We can construct a solution to the heat equation starting from an arbitrary initial point $u_0 \in H^1(M)$ as gradient flow of the energy functional $\mathcal{E}$ on $L^2(M; m)$. As $L^2(M; m)$ is a Hilbert space, we can certainly apply Crandall and Liggett's classical technique ([CL], see also [AGS]).

Theorem 5.4. ([OS]) *For each $u_0 \in H^1(M)$ and $T > 0$, there exists a unique solution u to the heat equation lying in $L^2([0,T], H^1(M)) \cap H^1([0,T], L^2(M))$. Moreover, for each $t \in (0,T)$, the distributional Laplacian Δu_t is absolutely continuous with respect to m and we have*

$$\lim_{\delta \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(u_{t+\delta})}{\delta} = \|\Delta u_t\|_{L^2}^2$$

for all $t > 0$.

Uniqueness in Theorem 5.4 is a consequence of the L^2-contraction property which asserts that, given two solutions u, v to the heat equation with $\int_M u\,dm = \int_M v\,dm$,

$$\|u_t - v_t\|_{L^2} \le e^{-t\kappa\lambda} \|u_0 - v_0\|_{L^2}$$

holds for all $t > 0$. Note that the L^2-contraction property follows from the $(\kappa\lambda)$-convexity of $\mathcal{E}$, for a solution to the heat equation is gradient flow of $\mathcal{E}$ in $L^2(M;m)$.

As for regularity, we can show that every solution to the heat equation admits a Hölder continuous version and it is H^2 in the space variable x as well as $C^{1,\alpha}$ in both variables t and x for some $\alpha > 0$. However, u is not C^2 in general. The typical (noncompact) example is a Gaussian kernel of a Minkowski space, which is C^2 at the origin if and only if the Minkowski space happens to be a Hilbert space.

5.3. The Finsler structure of the Wasserstein space

We next present another characterization of heat flow as gradient flow of the relative entropy in the *reverse* Wasserstein space (i.e., the Wasserstein space with respect to the reverse Finsler structure). To do so, we introduce a Finsler structure of the Wasserstein space in accordance with Otto's formulation in [Ot] (see also [AGS], [Lo2]).

Given arbitrary $\varphi \in C^2(M)$, the function $\varepsilon\varphi$ is $(d^2/2)$-convex for sufficiently small $\varepsilon > 0$. This observation (together with Theorem 3.1) leads us to the following notion of tangent space:

$$T_\mu \mathcal{P} := \overline{\{V = \nabla\varphi \,|\, \varphi \in C^\infty(M)\}},$$

where the closure is taken with respect to the Minkowski norm

$$F_W(\mu, V) := \left(\int_M F\big(V(x)\big)^2 \, d\mu(x) \right)^{1/2}.$$

Note the dependence of $F_W(\mu, \cdot)$ on the base point μ, which generates the nontrivial geometric structure of $(\mathcal{P}(M), F_W)$.

Every locally Lipschitz continuous curve $(\mu_t)_{t\in(0,1)}$ is accompanied with the unique Borel vector field $\Psi_t(x) = \Psi(t,x) \in T_xM$ on $(0,1)\times M$ with $F(\Psi) \in L^\infty_{\mathrm{loc}}((0,1)\times M; d\mu_t dt)$ satisfying the *continuity equation*

$$\partial_t \mu_t + \mathrm{div}(\Psi_t \mu_t) = 0$$

in the weak sense that

$$\int_0^1 \int_M \{\partial_t \phi_t + D\phi_t \cdot \Psi_t\}\, d\mu_t\, dt = 0$$

holds for all $\phi \in C_c^\infty((0,1)\times M)$. We consider Ψ as the tangent vector field of the curve $(\mu_t)_{t\in(0,1)}$ and denote it by $\dot{\mu}_t = \Psi_t \in T_{\mu_t}\mathcal{P}$.

Given $\mu, \nu \in \mathcal{P}(M)$, we have

$$d_2^W(\mu,\nu) = \inf_{(\mu_t)_{t\in[0,1]}} \left(\int_0^1 F_W(\mu_t, \dot{\mu}_t)\, dt \right)^{1/2},$$

where $(\mu_t)_{t\in[0,1]}$ runs over all Lipschitz continuous curves in $\mathcal{P}(M)$ with $\mu_0 = \mu$ and $\mu_1 = \nu$. Thus our Finsler structure realizes the Wasserstein distance.

5.4. Heat flow as gradient flow in the Wasserstein space

We define the cotangent space at given $\mu \in \mathcal{P}(M)$ in a similar manner to the tangent space as

$$T_\mu^*\mathcal{P} := \overline{\{\alpha = D\varphi \,|\, \varphi \in C^\infty(M)\}},$$
$$F_W^*(\mu,\alpha) := \left(\int_M F^*\big(\alpha(x)\big)^2\, d\mu(x) \right)^{1/2},$$

where the closure is taken with respect to $F_W^*(\mu,\cdot)$. Then the associated Legendre transform $\mathcal{L}_W^*(\mu,\cdot) : T_\mu^*\mathcal{P} \longrightarrow T_\mu\mathcal{P}$ at μ is defined in the pointwise way by $\mathcal{L}_W^*(\mu,\alpha)(x) := \mathcal{L}^*(\alpha(x))$. We also define the exponential map $\exp_\mu : T_\mu\mathcal{P} \longrightarrow \mathcal{P}(M)$ by $\exp_\mu(V) := (\exp V)_\sharp\mu$.

Given a function S on (a subset of) $\mathcal{P}(M)$, we say that S is differentiable at $\mu \in \mathcal{P}(M)$ if there is $\alpha \in T_\mu^*\mathcal{P}$ such that

$$\int_M \alpha \cdot V\, dm = \lim_{t\downarrow 0} \frac{S(\exp_\mu(tV)) - S(\mu)}{t}$$

holds for all $V = \nabla\varphi \in T_\mu\mathcal{P}$ with $\varphi \in C^\infty(M)$. Then we denote α by $DS(\mu)$ and call it the derivative of S at μ, and the gradient vector of S at μ is defined as $\nabla_W S(\mu) := \mathcal{L}_W^*(\mu, DS(\mu))$.

Definition 5.5. (Gradient flow) A continuous curve $(\mu_t)_{t \ge 0} \subset \mathcal{P}(M)$ which is locally Lipschitz continuous on $(0, \infty)$ is called a *gradient flow* of S if $\dot{\mu}_t = \nabla_W(-S)(\mu_t)$ holds at a.e. $t \in (0, \infty)$.

Note the difference between $\nabla_W(-S)$ and $-\nabla_W S$. The gradient vector of $-\mathrm{Ent}_m$ can be described explicitly.

Proposition 5.6. ([OS]) *Take absolutely continuous $\mu = \rho m \in \mathcal{P}(M)$ with $\rho \in H^1(M)$. If $-\log \rho \notin H^1(M; \mu)$, then $-\mathrm{Ent}_m$ is not differentiable at μ. If $-\log \rho \in H^1(M; \mu)$, then $-\mathrm{Ent}_m$ is differentiable at μ and the gradient vector is given by*

$$\nabla_W(-\mathrm{Ent}_m)(\mu) = \frac{1}{\rho}\nabla(-\rho) \in T_\mu \mathcal{P}.$$

In particular, its norm squared $F_W(\mu, \nabla_W(-\mathrm{Ent}_m)(\mu))^2$ coincides with the Fisher information *with respect to the reverse Finsler structure $\overleftarrow{F}$:*

$$\overleftarrow{I}(\mu) := \int_M \overleftarrow{F}\big(\overleftarrow{\nabla}\rho\big)^2 \frac{1}{\rho}\, dm = \int_M F\big(\nabla(-\rho)\big)^2 \frac{1}{\rho}\, dm.$$

We finally obtain a characterization of heat flow as gradient flow of the relative entropy in the reverse Wasserstein space. Compare this with the Euclidean and Riemannian cases in [JKO], [Oh1] and [Sa].

Theorem 5.7. ([OS]) *Let $(\mu_t)_{t \ge 0} \subset \mathcal{P}(M)$ be a continuous curve which is locally Lipschitz continuous on $(0, \infty)$, and assume that $\mu_t = \rho_t m$ with $\rho_t \in H^1(M)$ a.e. $t \in (0, \infty)$. Then $(\mu_t)_{t \ge 0}$ is a gradient flow of the relative entropy if and only if $(\rho_t)_{t \ge 0}$ is a heat flow with respect to the reverse Finsler structure $\overleftarrow{F}$ of F.*

References

[AT] J. C. Álvarez-Paiva and A. C. Thompson, Volumes in normed and Finsler spaces, In: A sampler of Riemann-Finsler geometry, 1–48, Math. Sci. Res. Inst. Publ., **50**, Cambridge Univ. Press, Cambridge, 2004.

[AGS] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures, Birkhäuser Verlag, Basel, 2005.

[BE] D. Bakry and M. Émery, Diffusions hypercontractives (French), Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., 1123, Springer, Berlin, 1985.

[BCS] D. Bao, S.-S. Chern and Z. Shen, An introduction to Riemann-Finsler geometry, Springer-Verlag, New York, 2000.

[Br] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. **44** (1991), 375–417.

[CC] J. Cheeger and T. H. Colding, On the structure of spaces with Ricci curvature bounded below. I, II, III, J. Differential Geom. **46** (1997), 406–480; ibid. **54** (2000), 13–35; ibid. **54** (2000), 37–74.

[CMS] D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschläger, A Riemannian interpolation inequality á la Borell, Brascamp and Lieb, Invent. Math. **146** (2001), 219–257.

[CL] M. G. Crandall and T. M. Liggett, Generation of semi-groups of nonlinear transformations on general Banach spaces, Amer. J. Math. **93** (1971), 265–298.

[Ga] R. J. Gardner, The Brunn-Minkowski inequality, Bull. Amer. Math. Soc. (N.S.) **39** (2002), 355–405.

[GM] M. Gromov and V. D. Milman, A topological application of the isoperimetric inequality, Amer. J. Math. **105** (1983), 843–854.

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

[KK] A. Kristály and L. Kozma, *Metric characterization of Berwald spaces of non-positive flag curvature*, J. Geom. Phys. **56** (2006), 1257–1270.

[KVK] A. Kristály, C. Varga and L. Kozma, *The dispersing of geodesics in Berwald spaces of non-positive flag curvature*, Houston J. Math. **30** (2004), 413–420.

[Le] M. Ledoux, The concentration of measure phenomenon, American Mathematical Society, Providence, RI, 2001.

[Lo1] J. Lott, Some geometric properties of the Bakry-Émery-Ricci tensor, Comment. Math. Helv. **78** (2003), 865–883.

[Lo2] J. Lott, Some geometric calculations on Wasserstein space, Comm. Math. Phys. **277** (2008), 423–437.

[LV1] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. **169** (2009), 903–991.

[LV2] J. Lott and C. Villani, Weak curvature conditions and functional inequalities, J. Funct. Anal. **245** (2007), 311–333.

[Mc] R. J. McCann, Polar factorization of maps on Riemannian manifolds, Geom. Funct. Anal. **11** (2001), 589–608.

[Oh1] S. Ohta, Gradient flows on Wasserstein spaces over compact Alexandrov spaces, Amer. J. Math. **131** (2009), 475–516.

[Oh2] S. Ohta, Uniform convexity and smoothness, and their applications in Finsler geometry, Math. Ann. **343** (2009), 669–699.

[Oh3] S. Ohta, Finsler interpolation inequalities, Calc. Var. Partial Differential Equations **36** (2009), 211–249.

[OS] S. Ohta and K.-T. Sturm, Heat flow on Finsler manifolds, Comm. Pure Appl. Math. **62** (2009), 1386–1433.

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

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

[Qi] Z. Qian, Estimates for weighted volumes and applications, Quart. J. Math. Oxford Ser. (2) **48** (1997), 235–242.

[vRS] M.-K. von Renesse and K.-T. Sturm, Transport inequalities, gradient estimates, entropy and Ricci curvature, Comm. Pure Appl. Math. **58** (2005), 1–18.

[Sa] G. Savaré, Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds, C. R. Math. Acad. Sci. Paris **345** (2007), 151–154.

[Sh1] Z. Shen, Volume comparison and its applications in Riemann-Finsler geometry, Adv. Math. **128** (1997), 306–328.

[Sh2] Z. Shen, Lectures on Finsler geometry, World Scientific Publishing Co., Singapore, 2001.

[St1] K.-T. Sturm, Convex functionals of probability measures and nonlinear diffusions on manifolds, J. Math. Pures Appl. **84** (2005), 149–168.

[St2] K.-T. Sturm, On the geometry of metric measure spaces, Acta Math. **196** (2006), 65–131.

[St3] K.-T. Sturm, On the geometry of metric measure spaces. II, Acta Math. **196** (2006), 133–177.

[Vi1] C. Villani, Topics in optimal transportation, American Mathematical Society, Providence, RI, 2003.

[Vi2] C. Villani, Optimal transport, old and new, Springer-Verlag, Berlin, 2009.

Department of Mathematics, Faculty of Science, Kyoto University, Kyoto 606-8502, JAPAN

E-mail address: sohta@math.kyoto-u.ac.jp