Uniform convexity and smoothness, and their applications in Finsler geometry*†

Shin-ichi OHTA‡
Department of Mathematics, Faculty of Science,
Kyoto University, Kyoto 606-8502, JAPAN
e-mail: sohta@math.kyoto-u.ac.jp

Abstract

We generalize the Alexandrov-Toponogov comparison theorems to Finsler manifolds. Under suitable upper (lower, resp.) bounds on the flag and tangent curvatures together with the 2-uniform convexity (smoothness, resp.) of tangent spaces, we show the 2-uniform convexity (smoothness, resp.) of Finsler manifolds. As applications, we prove the almost everywhere existence of the second order differentials of semi-convex functions and of c-concave functions with the quadratic cost function.

1 Introduction

The aim of this article is to present an intersection of comparison Finsler geometry and the geometry of Banach spaces. We refer to [BCS], [Eg], [Sh1], [Sh2], [Sh3] and [WX] for known results in comparison Finsler geometry. One of the milestones in comparison Riemannian geometry is the Alexandrov-Toponogov comparison theorems which have generated, beyond pure Riemannian geometry, the deep theories of CAT(κ)-spaces and Alexandrov spaces (see [BH] and [BBI]). However, it is well-known that a Finsler manifold can be a CAT(κ)-space or an Alexandrov space only if it is Riemannian. Similarly, a Banach space is a CAT(0)-space or an Alexandrov space of nonnegative curvature if and only if it happens to be a Hilbert space. Therefore these conditions are too restrictive in Finsler geometry, and we propose a generalization of them from the viewpoint of Banach space theory.

In a fascinating paper [BCL], Ball, Carlen and Lieb have introduced two important notions in Banach space theory, called the p-uniform convexity and the q-uniform smoothness. A Banach space $(X, |\cdot|)$ is said to be *p-uniformly convex* for $2 \le p < \infty$ if there is

*Mathematics Subject Classification (2000): 53C60, 46B20, 58C20.
†Keywords: Finsler geometry, Alexandrov-Toponogov comparison theorem, uniform convexity, uniform smoothness, convex function, c-concave function.
‡Partly supported by the JSPS fellowships for research abroad.

a constant $C \geq 1$ such that

$$\left|\frac{v+w}{2}\right|^p \leq \frac{1}{2}|v|^p + \frac{1}{2}|w|^p - C^{-p}\left|\frac{v-w}{2}\right|^p \tag{1.1}$$

holds for any $v, w \in X$. Similarly, the *q-uniform smoothness* for $1 < q \leq 2$ is defined by

$$\left|\frac{v+w}{2}\right|^q \geq \frac{1}{2}|v|^q + \frac{1}{2}|w|^q - S^q\left|\frac{v-w}{2}\right|^q. \tag{1.2}$$

Since then, these simple inequalities turned out to be useful instruments in Banach space theory and the geometry of Banach spaces. For instance, L_p-spaces are 2-uniformly convex with $C = 1/\sqrt{p-1}$ if $1 < p \leq 2$ ([BCL, Proposition 3]), and p-uniformly convex with $C = 1$ if $2 \leq p < \infty$ (Clarkson's inequality). By duality, L_q-spaces are 2-uniformly smooth with $S = \sqrt{q-1}$ if $2 \leq q < \infty$, and q-uniformy smooth with $S = 1$ if $1 < q \leq 2$. The quadratic cases $p = q = 2$ are of particular interest.

One natural generalization of the 2-uniform convexity (1.1) to a nonlinear metric space (X, d) is the following: For any point $x \in X$ and any minimal geodesic $\eta : [0,1] \longrightarrow X$, we have

$$d\left(x, \eta\left(\frac{1}{2}\right)\right)^2 \leq \frac{1}{2}d\big(x, \eta(0)\big)^2 + \frac{1}{2}d\big(x, \eta(1)\big)^2 - \frac{1}{4C^2}d\big(\eta(0), \eta(1)\big)^2. \tag{1.3}$$

In other words, $d^2/dt^2[d(x, \eta(t))^2] \geq 2C^{-2}d(\eta(0), \eta(1))^2$ in a weak sense. If $C = 1$, then the inequality (1.3) corresponds to the CAT(0)-property. Therefore the inequality (1.3) involves two different interpretations: a nonlinearized 2-uniform convexity as well as a weakened CAT(0)-property. (Although there is another possibility

$$d\left(x, \eta\left(\frac{1}{2}\right)\right)^2 \geq \frac{1}{2C^2}d\big(x, \eta(0)\big)^2 + \frac{1}{2}d\big(x, \eta(1)\big)^2 - \frac{1}{4}d\big(\eta(0), \eta(1)\big)^2,$$

we do not treat it in this article.) Similarly, the 2-uniform smoothness (1.2) is nonlinearized as

$$d\left(x, \eta\left(\frac{1}{2}\right)\right)^2 \geq \frac{1}{2}d\big(x, \eta(0)\big)^2 + \frac{1}{2}d\big(x, \eta(1)\big)^2 - \frac{S^2}{4}d\big(\eta(0), \eta(1)\big)^2, \tag{1.4}$$

and it contains Alexandrov spaces of nonnegative curvature as the special case of $S = 1$.

The inequalities (1.3) and (1.4) play interesting roles and have various applications in metric geometry, the nonlinearization of the geometry of Banach spaces and their related fields (see [Oh1], [Oh2], [Oh3], [Oh4] and [Oh5]). For instance, the 2-uniform smoothness finds its usefulness in Wasserstein geometry, as the L_2-Wasserstein space over a compact geodesic space X is 2-uniformly smooth if and only if so is X ([Oh4], [Sa]).

We will show that a Finsler manifold is 2-uniformly convex (smooth, resp.) if the flag curvature and Shen's tangent curvature are bounded above (below, resp.) and if tangent spaces are 2-uniformly convex (smooth, resp.) with a uniform bound on the constant C in (1.1) (S in (1.2), resp.). See Theorems 4.2, 5.1, Corollaries 4.4 and 5.2 for precise statements. These naturally extend the Alexandrov-Toponogov comparison theorems.

We remark that Theorem 5.1 (the local 2-uniform convexity) has already been observed by Shen [Sh3].

We provide two analytic applications. The first one (Theorem 6.6) is the Finsler analogue of the famous Alexandrov-Bangert theorem ([Al], [Ba]) on the almost everywhere existence of the second order differentials of (semi-)convex functions. The 2-uniform convexity appears to extend the theorem from convex functions to semi-convex functions. In the middle of the proof, we generalize Greene and Wu's approximation technique (Proposition 6.7), and it seems to be of independent interest.

The second application (Theorem 7.4) is the almost everywhere second order differentiability of c-concave functions with the quadratic cost function $c(x, y) = d(x, y)^2/2$. It follows from the above first application and the 2-uniform smoothness. The class of c-concave functions is known to play quite an important role in mass transport theory (see [Br], [CMS], [Mc], [RR], [Vi1] and [Vi2]). In fact, in a recent work of the author [Oh6], Theorem 7.4 acts as one of key ingredients of extending Lott, Sturm and Villani's remarkable theory ([LV1], [LV2], [St1], [St2]) to Finsler manifolds. To be precise, just like the Riemannian case, a kind of lower Ricci curvature bound turns out equivalent to the curvature-dimension condition. Then there are many applications in functional inequalities, interpolation inequalities and the concentration of measure phenomenon, and most of them are new in the Finsler setting.

The article is organized as follows: In Sections 2 and 3, we review the basics of Finsler geometry. Sections 4 and 5 are devoted to the 2-uniform smoothness and convexity of Finsler manifolds, respectively. Then we apply them to establishing the almost everywhere second order differentiability of semi-convex and c-concave functions in Sections 6 and 7, respectively.

Throughout the article, without otherwise indicated, Finsler manifolds are only positively homogeneous of degree one, and every geodesic has a constant speed (i.e., it is parametrized proportionally to the arclength).

Acknowledgements. I would like to thank Karl-Theodor Sturm for discussions from which this work started. This work was done while I was visiting Institut für Angewandte Mathematik, Universität Bonn. I am also grateful to the institute for its hospitality.

2 Preliminaries for Finsler geometry

In this and the next sections, we review a fundamental knowledge of Finsler geometry. We refer to [BCS] and [Sh3] as comprehensive references.

2.1 Fundamental and Cartan tensors

Let M be a connected C^∞-manifold. For $x \in M$, we denote by T_xM the tangent space at x, put $TM := \bigcup_{x \in M} T_xM$ and let $\pi : TM \longrightarrow M$ be 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 $(x^i, y^i)_{i=1}^n$ a local coordinate system on $\pi^{-1}(U) \subset TM$ given by, for $v \in \pi^{-1}(U)$,

$$v = \sum_i v^i \frac{\partial}{\partial x^i}\Big|_{\pi(v)} = \sum_i y^i(v) \frac{\partial}{\partial x^i}\Big|_{\pi(v)}.$$

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

(1) 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 $(x^i)_{i=1}^n : U \longrightarrow \mathbb{R}^n$ on $U \subset M$, the $n \times n$ matrix

$$g_{ij}(v) := \frac{1}{2} \frac{\partial^2 (F^2)}{\partial y^i \partial y^j}(v)$$

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

In other words, the function F provides a *Minkowski norm* on each tangent space, and is C^∞ in the horizontal direction. As a consequence, F is zero on the zero section 0 and positive on $TM \setminus 0$. Thanks to the positive homogeneity, $g_{ij}(\lambda v) = g_{ij}(v)$ holds for any $v \in TM \setminus 0$ and $\lambda > 0$. We emphasis that F is not necessarily *absolutely homogeneous*, that is, $F(v) \neq F(-v)$ may happen. This unreversibility of F causes the possible nonsymmetricity of the associated distance function.

The strong convexity certainly guarantees that F is strictly convex on each tangent space T_xM in the sense that $F(v + w) < F(v) + F(w)$ holds unless v and w are colinear. Moreover, the matrix $(g_{ij}(v))_{i,j=1}^n$ defines a Hilbertian (Riemannian) structure of $T_{\pi(v)}M$ and we denote it by g_v, i.e.,

$$g_v\bigg(\sum_i w_1^i \frac{\partial}{\partial x^i}\Big|_{\pi(v)}, \sum_j w_2^j \frac{\partial}{\partial x^j}\Big|_{\pi(v)} \bigg) := \sum_{i,j} g_{ij}(v) w_1^i w_2^j.$$

Note that $g_v(v, v) = F(v)^2$ follows from Euler's theorem (cf. [BCS, Theorem 1.2.1]). We call $g_{ij}(v)$ the *fundamental tensor* and further define the *Cartan tensor* by, for $v \in TM \setminus 0$,

$$A_{ijk}(v) := \frac{F(v)}{2} \frac{\partial g_{ij}}{\partial y^k}(v) = \frac{F(v)}{4} \frac{\partial^3 (F^2)}{\partial y^i \partial y^j \partial y^k}(v).$$

Observe that $A_{ijk}(\lambda v) = A_{ijk}(v)$ holds for any $\lambda > 0$. If F is coming from a Riemannian structure, then g_v coincides with the original Riemannian structure for all $v \in TM \setminus 0$, therefore the Cartan tensors vanish everywhere. In fact, the converse is also true, namely F is Riemannian if and only if the Cartan tensors vanish everywhere on $TM \setminus 0$.

2.2 Chern connection and covariant derivatives

We define the *formal Christoffel symbol* by

$$\gamma^i{}_{jk}(v) := \frac{1}{2}\sum_l g^{il}(v)\left\{\frac{\partial g_{lj}}{\partial x^k}(v) - \frac{\partial g_{jk}}{\partial x^l}(v) + \frac{\partial g_{kl}}{\partial x^j}(v)\right\}$$

for $v \in TM \setminus 0$, and also define

$$N^i{}_j(v) := \sum_k \gamma^i{}_{jk}(v)v^k - \frac{1}{F(v)}\sum_{k,l,m} A^i{}_{jk}(v)\gamma^k{}_{lm}(v)v^l v^m,$$

where (g^{ij}) stands for the inverse matrix of (g_{ij}) and $A^i{}_{jk} := \sum_l g^{il}A_{ljk}$.

Given a connection ∇ on the pulled-back tangent bundle π^*TM, we denote its connection one-forms by $\omega_j{}^i$, that is,

$$\nabla_v \frac{\partial}{\partial x^j} = \sum_i \omega_j{}^i(v)\frac{\partial}{\partial x^i}, \qquad \nabla_v dx^i = -\sum_j \omega_j{}^i(v)\, dx^j.$$

Different from the Riemannian situation, there are several connections (due to Cartan, Chern, Berwald etc.) each of which is canonical in its own way. We use only one of them in this article.

Definition 2.2 (Chern connection) Let (M, F) be a C^∞-Finsler manifold. Then there is a unique connection ∇ on the pulled-back tangent bundle π^*TM, called the *Chern connection*, whose connection one-forms $\omega_j{}^i$ satisfy the following conditions: Fix a local coordinate system $(x^i)_{i=1}^n : U \longrightarrow \mathbb{R}^n$ on $U \subset M$.

(1) (Torsion-freeness) For any $i = 1, 2, \ldots, n$, we have

$$\sum_j dx^j \wedge \omega_j{}^i = 0.$$

(2) (Almost g-compatibility) For any $i, j = 1, 2, \ldots, n$, we have

$$dg_{ij} - \sum_k (g_{kj}\omega_i{}^k + g_{ik}\omega_j{}^k) = \frac{2}{F}\sum_k A_{ijk}\,\delta y^k,$$

where we set $\delta y^k := dy^k + \sum_l N^k{}_l\, dx^l$.

Henceforce, ∇ always stands for the Chern connection on π^*TM. The torsion-freeness says that the connection one-form $\omega_j{}^i$ does not have any dy^k-term, so that we can write $\omega_j{}^i = \sum_k \Gamma^i{}_{jk}\, dx^k$. The torsion-freeness also implies $\Gamma^i{}_{jk} = \Gamma^i{}_{kj}$ and, together with the almost g-compatibility, we find the explicit formula

$$\Gamma^i{}_{jk} = \gamma^i{}_{jk} - \frac{1}{F}\sum_{l,m} g^{il}(A_{ljm}N^m{}_k - A_{jkm}N^m{}_l + A_{klm}N^m{}_j).$$

If (M,F) is Riemannian, then the Cartan tensors vanish everywhere, and hence the almost g-compatibility reduces to the *g-compatibility* $dg_{ij} = \sum_k (g_{kj}\omega_i{}^k + g_{ik}\omega_j{}^k)$, therefore the Chern connection is nothing but the Levi-Civita connection. We say that a Finsler manifold (M,F) is of *Berwald type* if $\Gamma^i{}_{jk}(v)$ depends only on $x = \pi(v)$ (i.e., $\Gamma^i{}_{jk}$ is fiber-wise constant). For instance, Riemannian manifolds and Minkowski spaces are of Berwald type. Roughly speaking, a Finsler manifold of Berwald type is modeled on a single Minkowski space. Finsler manifolds of Berwald type have already provided a rich family of non-Riemannian spaces.

For a C^∞-vector field X on M and two nonzero vectors $v, w \in T_xM \setminus 0$, we define the *covariant derivative* $D_v^w X$ with *reference vector* w as

$$(D_v^w X)(x) := \sum_{i,j} \left\{ v^j \frac{\partial X^i}{\partial x^j}(x) + \sum_k \Gamma^i{}_{jk}(w) v^j X^k(x) \right\} \frac{\partial}{\partial x^i}\Big|_x, \tag{2.1}$$

where $X(x) = \sum_i X^i(x)(\partial/\partial x^i)|_x$. We usually choose $w = v$ or $X(x)$.

2.3 Flag and tangent curvatures

The Chern connection ∇ leads the corresponding curvature two-form

$$\Omega_j{}^i(v) := d\omega_j{}^i - \sum_k \omega_j{}^k \wedge \omega_k{}^i,$$

where we put $\omega \wedge \tau := \omega \otimes \tau - \tau \otimes \omega$. It can be rewritten as

$$\Omega_j{}^i(v) = \frac{1}{2} \sum_{k,l} R_j{}^i{}_{kl}(v)\, dx^k \wedge dx^l + \frac{1}{F(v)} \sum_{k,l} P_j{}^i{}_{kl}(v)\, dx^k \wedge \delta y^l,$$

where we impose $R_j{}^i{}_{kl} = -R_j{}^i{}_{lk}$. Recall that $\delta y^l = dy^l + \sum_m N^l{}_m\, dx^m$. Here the $(\delta y^k \wedge \delta y^l)$-terms do not appear by virtue of the torsion-freeness of the Chern connection.

Given two linearly independent vectors $v, w \in T_xM \setminus 0$, we define the *flag curvature* by

$$\mathcal{K}(v,w) := \frac{g_v(R^v(w,v)v, w)}{g_v(v,v)g_v(w,w) - g_v(v,w)^2},$$

where we set, for $v = \sum_i v^i(\partial/\partial x^i)|_x$ and $w = \sum_i w^i(\partial/\partial x^i)|_x$,

$$R^v(w,v)v := \sum_{i,j,k,l} v^j R_j{}^i{}_{kl}(v) w^k v^l \frac{\partial}{\partial x^i}\Big|_x. \tag{2.2}$$

Observe that $\mathcal{K}(\lambda v, \mu w) = \mathcal{K}(v,w)$ holds for any $\lambda > 0$ and $\mu \neq 0$. Unlike the Riemannian case, the flag curvature $\mathcal{K}(v,w)$ depends not only on the *flag* $\{\lambda v + \mu w \,|\, \lambda, \mu \in \mathbb{R}\}$, but also on the *flag pole* $\{\lambda v \,|\, \lambda > 0\}$. One merit of the flag curvature is its independence of a choice of connections.

We next recall another kind of curvature introduced by Shen (see [Sh2] and [Sh3]). For two nonzero vectors $v, w \in T_xM \setminus 0$, we define the *tangent curvature* by

$$\mathcal{T}_v(w) := \sum_{i,j,k,l} v^l g_{li}(v) \{ \Gamma^i{}_{jk}(w) - \Gamma^i{}_{jk}(v) \} w^j w^k.$$

If V and W are C^∞-vector fields on M, then $\mathcal{T}_V(W)$ is rewritten as, by using the covariant derivative (2.1),

$$\mathcal{T}_V(W) = g_V(D_W^W W - D_W^V W, V). \tag{2.3}$$

Observe that $\mathcal{T}_v(v) = 0$ and $\mathcal{T}_{\lambda v}(\mu w) = \lambda \mu^2 \mathcal{T}_v(w)$ for any $\lambda, \mu > 0$. Hence we say that $\mathcal{T} \ge -\delta$ or $\mathcal{T} \le \delta$ for a nonnegative constant $\delta \ge 0$ if we have

$$\mathcal{T}_v(w) \ge -\delta F(v) F(w)^2 \quad \text{or} \quad \mathcal{T}_v(w) \le \delta F(v) F(w)^2 \tag{2.4}$$

for all $v, w \in T_xM \setminus 0$ and $x \in M$, respectively. Although somewhat stronger bounds

$$\mathcal{T}_v(w) \ge -\delta F(v) \{ g_v(w,w) - g_v\big(w, v/F(v)\big)^2 \} \; \big[\ge -\delta \mathcal{S}(x)^2 F(v) F(w)^2 \big],$$
$$\mathcal{T}_v(w) \le \delta F(v) \{ g_v(w,w) - g_v\big(w, v/F(v)\big)^2 \} \; \big[\le \delta \mathcal{S}(x)^2 F(v) F(w)^2 \big]$$

as in [Sh2] and [Sh3] might be more natural in some situations, we prefer (2.4) because they are suitable to describe our results. (See (4.1) for the definition of $\mathcal{S}(x)$.) It is known that the tangent curvature $\mathcal{T}$ vanishes everywhere if and only if (M, F) is of Berwald type (see [Sh2, Proposition 3.1], [Sh3, Proposition 10.1.1] and also [BCS, Proposition 10.2.1]), so that the tangent curvature $\mathcal{T}$ measures the variation of tangent Minkowski spaces in a sense.

3 Variational formulas and Jacobi fields

3.1 Geodesics and exponential map

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

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

Then the corresponding distance function $d : M \times M \longrightarrow [0, \infty)$ is given by $d(x_1, x_2) := \inf_\eta L(\eta)$, where the infimum is taken over all C^∞-curves η from x_1 to x_2. We emphasis that d is not necessarily symmetric (i.e., $d(x_1, x_2) \neq d(x_2, x_1)$ may happen) because F is merely positively homogeneous. Nonetheless, d is positive away from the diagonal set and satisfies the triangle inequality $d(x_1, x_3) \le d(x_1, x_2) + d(x_2, x_3)$. We define the *forward* and *backward open balls* of center $x \in M$ and radius $r > 0$ by

$$B^+(x, r) := \{ z \in M \,|\, d(x, z) < r \}, \qquad B^-(x, r) := \{ z \in M \,|\, d(z, x) < r \}.$$

Similarly, *forward* and *backward closed balls* are denoted by $\bar B^+(x, r)$ and $\bar B^-(x, r)$, respectively.

A C^∞-curve $\eta : [0,r] \longrightarrow M$ is called a *geodesic* if it satisfies $D_{\dot\eta}^{\dot\eta}[\dot\eta/F(\dot\eta)] = 0$ on $[0,r]$ (compare this with the first variation formula (3.3) below). We remark that the reverse curve $\bar\eta(t) := \eta(r-t)$ may not be a geodesic. Let V and W be C^∞-vector fields along a geodesic $\eta : [0,r] \longrightarrow M$. Then it holds that (cf. [BCS, Exercise 5.2.3])

$$\frac{d}{dt}\Big[g_{\dot\eta(t)}\big(V(t),W(t)\big)\Big] = g_{\dot\eta(t)}\big(D_{\dot\eta}^{\dot\eta}V(t),W(t)\big) + g_{\dot\eta(t)}\big(V(t),D_{\dot\eta}^{\dot\eta}W(t)\big). \tag{3.1}$$

Note that a C^∞-curve $\eta : [0,r] \longrightarrow M$ is a geodesic of constant speed (i.e., $F(\dot\eta)$ is constant) if and only if we have $D_{\dot\eta}^{\dot\eta}\dot\eta \equiv 0$. As was indicated at the end of the introduction, every geodesic in this article will have constant speed.

We define the *exponential map* by $\exp v = \exp_{\pi(v)} v := \eta(1)$ for $v \in TM$ if there is a geodesic $\eta : [0,1] \longrightarrow M$ with $\dot\eta(0) = v$. The exponential map is only C^1 at the zero section, and is C^2 at the zero section if and only if (M,F) is of Berwald type. Moreover, the squared distance function $d(x,\cdot)^2$ from a point $x \in M$ is C^2 at x for all $x \in M$ if and only if (M,F) is Riemannian ([Sh1, Proposition 2.2]). The lack of C^2-smoothness is troublesome from the analytic viewpoint, and we indeed need some tricks in Section 6 which are unnecessary in the Riemannian case. Nevertheless, the following useful lemma holds true. We remark that $\xi_v(t) = \exp tv$ holds only for $t \ge 0$ in the lemma.

Lemma 3.1 (cf. [BCS, page 125]) *Let (M,F) be a C^∞-Finsler manifold and take an open set $U \subset M$ whose closure $\overline{U}$ is compact. Then there exists a positive constant $\varepsilon > 0$ such that the map*

$$\{v \in \pi^{-1}(U) \,|\, 0 < F(v) < \varepsilon\} \times (-1,1) \ni (v,t) \longmapsto \xi_v(t) \in M$$

is well-defined and C^∞, where $\xi_v : (-1,1) \longrightarrow M$ is the geodesic with $\dot\xi_v(0) = v$.

A C^∞-Finsler manifold (M,F) is said to be *forward geodesically complete* if the exponential map is defined on the entire TM, in other words, if there is a geodesic $\eta : [0,\infty) \longrightarrow M$ with $\dot\eta(0) = v$ for any given $v \in TM$. Then any two points in M can be connected by a minimal geodesic, i.e., a geodesic whose arclength coincides with the distance from the initial point to the terminal point.

Given a point $x \in M$, the *injectivity radius* $\mathrm{inj}(x)$ at x is the supremum of $r > 0$ such that the geodesic $\eta(t) := \exp_x tv$ exists and is minimal on $[0,1]$ for all $v \in T_xM$ with $F(v) \le r$. Then $\mathrm{inj}(x)$ is positive and possibly infinite, and the function $x \longmapsto \mathrm{inj}(x)$ is continuous (cf. [BCS, Proposition 8.4.1(2)]). The exponential map at x is a C^1-diffeomorphism from $B^+(0,\mathrm{inj}(x)) = \{v \in T_xM \,|\, F(v) < \mathrm{inj}(x)\}$ to $B^+(x,\mathrm{inj}(x))$, and is C^∞ on $B^+(0,\mathrm{inj}(x)) \setminus 0$.

3.2 Gradient vectors and Hessians

We denote by (T_x^*M, F^*) the dual space of (T_xM, F). 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 $df_x \in T_x^*M$. That is to say, $\mathrm{grad} f(x) := \mathfrak{L}_x(df_x) \in T_xM$ is the unique vector satisfying $F(\mathrm{grad} f(x)) = F^*(df_x)$ and $df_x(\mathrm{grad} f(x)) = F^*(df_x)^2$. Note that $\mathrm{grad} f(x)$ faces toward the steepest direction of f.

In the Riemannian situation, the Hessian of a C^2-function $f : M \longrightarrow \mathbb{R}$ at $x \in M$ is a bilinear function $\mathrm{Hess}\, f : T_xM \times T_xM \longrightarrow \mathbb{R}$ defined by $\mathrm{Hess}\, f(v,w) := g(\nabla_v(\mathrm{grad}\, f), w)$. In particular, for a geodesic $\eta(t) := \exp_x tv$, we observe

$$\mathrm{Hess}\, f(v,v) = \frac{d}{dt}\Big|_{t=0}\Big[g\big(\mathrm{grad}\, f \circ \eta(t), \dot\eta(t)\big)\Big] = \frac{d^2}{dt^2}\Big|_{t=0}\big[f \circ \eta(t)\big].$$

In the Finsler case, the Hessian can have only the second, restricted meaning. We define the *Hessian* of a C^2-function $f : M \longrightarrow \mathbb{R}$ on a Finsler manifold (M,F) by

$$\mathrm{Hess}\, f(v) := \frac{d^2}{dt^2}\Big|_{t=0}\big[f(\exp_x tv)\big], \tag{3.2}$$

where $v \in T_xM$.

3.3 Variational formulas for arclength

We consider a C^∞-variation $\sigma : [0,r] \times (-\varepsilon,\varepsilon) \longrightarrow M$ and define

$$T(t,s) := \frac{\partial \sigma}{\partial t}(t,s), \qquad U(t,s) := \frac{\partial \sigma}{\partial s}(t,s).$$

Then the first variation of arclength is written as (cf. [BCS, Exercise 5.2.4])

$$\frac{\partial}{\partial s} L\big(\sigma(\cdot,s)\big) = \frac{g_{T(r)}(U,T)}{F(T(r))} - \frac{g_{T(0)}(U,T)}{F(T(0))} - \int_0^r g_{T(t)}\left(U, D_T^T\left[\frac{T}{F(T)}\right]\right) dt. \tag{3.3}$$

We abbreviated as $T(t) := T(t,s)$, $g_{T(r)}(U,T) := g_{T(r,s)}(U(r,s),T(r,s))$ and so on in the right-hand side for brevity.

For a C^∞-variation $\sigma : [0,r] \times (-\varepsilon,\varepsilon) \longrightarrow M$ such that $\sigma(\cdot,0)$ is a geodesic, its second variation of arclength is calculated that (cf. [BCS, Exercise 5.2.7])

$$\begin{aligned}
&\frac{\partial^2}{\partial s^2} L\big(\sigma(\cdot,0)\big) \\
&= \int_0^r \frac{1}{F(T(t))}\big\{g_{T(t)}(D_T^T U, D_T^T U) - g_{T(t)}(R^T(U,T)T,U)\big\}\, dt \\
&\quad + \frac{g_{T(r)}(D_U^T U, T)}{F(T(r))} - \frac{g_{T(0)}(D_U^T U, T)}{F(T(0))} - \int_0^r \frac{1}{F(T(t))}\left\{\frac{\partial F(T)}{\partial s}(t,0)\right\}^2 dt.
\end{aligned} \tag{3.4}$$

We again abbreviated as $T(t) := T(t,0)$ etc. Recall (2.2) for the definition of $R^T(U,T)T$.

3.4 Jacobi fields and Rauch's comparison theorem

A C^∞-vector field J along a geodesic $\eta : [0,r] \longrightarrow M$ is called a *Jacobi field* if it satisfies

$$D_{\dot\eta}^{\dot\eta} D_{\dot\eta}^{\dot\eta} J + R^{\dot\eta}(J,\dot\eta)\dot\eta = 0$$

on $[0,r]$. Any Jacobi field is represented as the variational vector field of a geodesic variation and vice versa (cf. [BCS, Section 5.4, Exercise 7.1.1]). The Finsler version of

Rauch's comparison theorem plays an essential role in Section 5. Here we prove it for thoroughness. For $k \in \mathbb{R}$, we define the function $\mathbf{s}_k$ by

$$\mathbf{s}_k(t) := \begin{cases} (1/\sqrt{k})\sin(\sqrt{k}t) & \text{for } t \in [0, \pi/\sqrt{k}] \quad \text{if } k > 0, \\ t & \text{for } t \in [0, \infty) \quad \text{if } k = 0, \\ (1/\sqrt{-k})\sinh(\sqrt{-k}t) & \text{for } t \in [0, \infty) \quad \text{if } k < 0. \end{cases}$$

In other words, $\mathbf{s}_k$ is the solution of the differential equation $f'' + kf = 0$ with the initial conditions $f(0) = 0$ and $f'(0) = 1$.

Theorem 3.2 (Rauch's comparison theorem, cf. [BCS, Corollary 9.8.1]) *Let (M, F) be a C^∞-Finsler manifold and J be a Jacobi field along a unit speed geodesic $\eta : [0, r] \longrightarrow M$ with $J(0) = 0$. Assume $\mathcal{K} \le k$ for some $k \ge 0$ and $r < \pi/\sqrt{k}$ if $k > 0$. Then we have*

$$\frac{1}{2}\frac{d}{dt}\Big|_{t=r}\big[g_{\dot\eta(t)}(J, J)\big] \ge \frac{\mathbf{s}_k'(r)}{\mathbf{s}_k(r)} g_{\dot\eta(r)}(J, J).$$

Proof. Throughout the proof, we put $g_t := g_{\dot\eta(t)}$ and $J' := D_{\dot\eta}^{\dot\eta} J$ for simplicity. If $J(r) = 0$, then the assertion is clear, so that we suppose $J(r) \neq 0$. We may also assume $J \neq 0$ on $(0, r)$. Otherwise we replace J with $J|_{[\tau,r]}$ such that $J(\tau) = 0$ and $J \neq 0$ on $(\tau, r]$, and observe that $\mathbf{s}_k'(r - \tau)/\mathbf{s}_k(r - \tau) \ge \mathbf{s}_k'(r)/\mathbf{s}_k(r)$.

It follows from (3.1) and the assumption $\mathcal{K} \le k$ that, at each $t \in (0, r)$,

$$\begin{aligned} \frac{d^2}{dt^2}\big[g_t(J, J)^{1/2}\big] &= \frac{d}{dt}\left[\frac{g_t(J, J')}{g_t(J, J)^{1/2}}\right] = \frac{g_t(J, J'') + g_t(J', J')}{g_t(J, J)^{1/2}} - \frac{g_t(J, J')^2}{g_t(J, J)^{3/2}} \\ &= -\frac{g_t(R^{\dot\eta}(J, \dot\eta)\dot\eta, J)}{g_t(J, J)^{1/2}} + \frac{g_t(J', J')g_t(J, J) - g_t(J, J')^2}{g_t(J, J)^{3/2}} \\ &\ge -k g_t(J, J)^{-1/2}\{g_t(J, J)g_t(\dot\eta, \dot\eta) - g_t(J, \dot\eta)^2\} \\ &\ge -k g_t(J, J)^{1/2}. \end{aligned}$$

It implies

$$\frac{d}{dt}\left[\frac{d}{dt}\big[g_t(J, J)^{1/2}\big]\mathbf{s}_k(t) - g_t(J, J)^{1/2}\mathbf{s}_k'(t)\right] \ge 0.$$

This together with $\mathbf{s}_k(0) = 0$ and $J(0) = 0$ shows the required inequality. $\Box$

We remark that Rauch's comparison theorem is usually stated for all $k \in \mathbb{R}$ provided that J is $g_{\dot\eta}$-orthogonal to $\dot\eta$. However, the above statement (valid only for nonnegative k) is more convenient in our usage.

4 Uniform smoothness

In this section, we show the 2-uniform smoothness of Finsler manifolds.

4.1 Uniform smoothness and convexity constants

We introduce two quantities which estimate how far F is from Riemannian structures. For $x \in M$, we define

$$\mathcal{S}(x) := \sup_{v,w \in T_xM \setminus 0} \frac{g_w(v,v)^{1/2}}{g_v(v,v)^{1/2}} = \sup_{v,w \in T_xM \setminus 0} \frac{g_w(v,v)^{1/2}}{F(v)}, \tag{4.1}$$

$$\mathcal{C}(x) := \sup_{v,w \in T_xM \setminus 0} \frac{g_v(v,v)^{1/2}}{g_w(v,v)^{1/2}} = \sup_{v,w \in T_xM \setminus 0} \frac{F(v)}{g_w(v,v)^{1/2}}. \tag{4.2}$$

Clearly we see $\mathcal{S}(x), \mathcal{C}(x) \geq 1$, and it will turn out that $\mathcal{S}(x)$ and $\mathcal{C}(x)$ measure the smoothness and convexity of F on T_xM, respectively (see Propositions 4.1 and 4.6 below). The *uniformity constant* in [Eg] amounts to $\sup_{x \in M} \max\{\mathcal{S}(x), \mathcal{C}(x)\}$ in this context.

Proposition 4.1 *Let (M,F) be a C^∞-Finsler manifold. Then the following three conditions are equivalent*:

(i) *We have $\mathcal{S}(x) = 1$ for all $x \in M$.*

(ii) *We have $\mathcal{C}(x) = 1$ for all $x \in M$.*

(iii) *(M,F) is a Riemannian manifold.*

Proof. If (M,F) is Riemannian, then clearly $\mathcal{S}(x) = \mathcal{C}(x) = 1$ holds for any $x \in M$. If we have $\mathcal{S}(x) = 1$ for any $x \in M$, then we find

$$F\left(\frac{v+w}{2}\right)^2 \geq \frac{1}{2}F(v)^2 + \frac{1}{2}F(w)^2 - \frac{1}{4}F(w-v)^2$$

for all $v, w \in T_xM \setminus 0$ by Proposition 4.6 below. We also find the reverse inequality by exchanging (v,w) for $((v+w)/2, (w-v)/2)$, so that equality holds. It implies $\mathcal{S}(x) = \mathcal{C}(x) = 1$, and hence $g_w(v,v) = F(v)^2$ for all $v, w \in T_xM \setminus 0$. Therefore g_w depends only on $x = \pi(w)$, and thus (M,F) is Riemannian. The case of $\mathcal{C} \equiv 1$ is similar. $\square$

4.2 Uniform smoothness of Finsler manifolds

Theorem 4.2 (Uniform smoothness) *Let (M,F) be a connected, forward geodesically complete C^∞-Finsler manifold. Assume that there are constants $k, \delta \geq 0$ and $S \geq 1$ for which*

$$\mathcal{K} \geq -k, \qquad \mathcal{T} \geq -\delta, \qquad \mathcal{S} \leq S.$$

Then we have, for any $x, z \in M$ and $v \in T_zM$ with $F(v) = 1$,

$$\begin{aligned} &\limsup_{s \to 0} \frac{1}{2s^2}\left\{d\big(x, \xi_v(-s)\big)^2 + d\big(x, \xi_v(s)\big)^2 - 2d(x,z)^2\right\} \\ &\leq \begin{cases} S^2 \dfrac{\sqrt{k}r\cosh(\sqrt{k}r)}{\sinh(\sqrt{k}r)} + r\delta & \text{if } k > 0, \\ S^2 + r\delta & \text{if } k = 0, \end{cases} \end{aligned}$$

where $\xi_v : (-\varepsilon, \varepsilon) \longrightarrow M$ is the geodesic with $\dot{\xi}_v(0) = v$ and $r := d(x,z)$.

Proof. If $z = x$, then we immediately obtain

$$\lim_{s\to 0} \frac{1}{2s^2}\big\{d\big(z,\xi_v(-s)\big)^2 + d\big(z,\xi_v(s)\big)^2\big\} = \frac{1}{2}\big\{F(-v)^2 + F(v)^2\big\}$$
$$= \frac{1}{2}\big\{g_{-v}(-v,-v)+1\big\} = \frac{1}{2}\big\{g_{-v}(v,v)+1\big\} \le S^2,$$

so that we assume $z \ne x$. Let us also assume $k > 0$. Then the case of $k = 0$ follows from taking the limit as k tends to zero. Fix a unit speed, minimal geodesic $\eta : [0,r] \longrightarrow M$ from x to z, and take a parallel vector field V along η (i.e., $D^{\dot\eta}_{\dot\eta} V \equiv 0$) such that $V(r) = v$. We also define the function $f(t) := \sinh(\sqrt{k}t)/\sinh(\sqrt{k}r)$. For sufficiently small $\varepsilon > 0$, we consider the variation $\sigma : [0,r] \times (-\varepsilon,\varepsilon) \longrightarrow M$ given by $\sigma(t,s) := \xi_{V(t)}(sf(t))$. Note that σ is C^∞ by Lemma 3.1. As in Section 3, we put

$$T(t,s) := \frac{\partial\sigma}{\partial t}(t,s), \qquad U(t,s) := \frac{\partial\sigma}{\partial s}(t,s),$$
$$T(t) := T(t,0) = \dot\eta(t), \qquad U(t) := U(t,0) = f(t)V(t).$$

As V is parallel along η, it holds that $(D^T_T U)(t) = f'(t)V(t)$. Thus the second variation formula (3.4) yields that

$$\frac{\partial^2}{\partial s^2} L\big(\sigma(\cdot,0)\big) = \int_0^r \Big[f'(t)^2 g_{T(t)}(V,V)$$
$$- \mathcal{K}\big(T(t),U(t)\big)\big\{f(t)^2 g_{T(t)}(V,V) - f(t)^2 g_{T(t)}(T,V)^2\big\}\Big]\,dt$$
$$+ g_{T(r)}(D^T_U U,T) - \int_0^r \bigg\{\frac{\partial(F(T))}{\partial s}(t,0)\bigg\}^2\,dt.$$

We deduce from (3.1), (4.1), (2.3) and (2.4) that

$$g_{T(t)}(V,V) = g_{T(r)}(v,v) \le \mathcal{S}(z)^2 F(v)^2 \le S^2,$$
$$g_{T(r)}(D^T_U U,T) = g_{T(r)}(D^U_U U,T) - \mathcal{T}_{T(r)}(v) \le \delta F(v)^2 = \delta,$$
$$\int_0^r \bigg\{\frac{\partial(F(T))}{\partial s}(t,0)\bigg\}^2 dt \ge \frac{1}{r}\bigg\{\int_0^r \frac{\partial(F(T))}{\partial s}(t,0)\,dt\bigg\}^2 = \frac{1}{r}\bigg\{\frac{\partial}{\partial s}L\big(\sigma(\cdot,0)\big)\bigg\}^2.$$

These together imply

$$\frac{\partial^2}{\partial s^2}\big[L\big(\sigma(\cdot,0)\big)^2\big] = 2r\frac{\partial^2}{\partial s^2}L\big(\sigma(\cdot,0)\big) + 2\bigg\{\frac{\partial}{\partial s}L\big(\sigma(\cdot,0)\big)\bigg\}^2$$
$$\le 2r\int_0^r \Big[f'(t)^2 S^2 + kf(t)^2\big\{g_{T(t)}(V,V) - g_{T(t)}(T,V)^2\big\}\Big]\,dt + 2r\delta$$
$$\le 2r\int_0^r \big\{f'(t)^2 + kf(t)^2\big\}S^2\,dt + 2r\delta$$
$$= 2r\bigg\{\frac{S^2 k}{\sinh^2(\sqrt{k}r)}\int_0^r \cosh(2\sqrt{k}t)\,dt + \delta\bigg\}$$
$$= 2r\bigg\{\frac{S^2\sqrt{k}\cosh(\sqrt{k}r)}{\sinh(\sqrt{k}r)} + \delta\bigg\}.$$

This completes the proof since

$$\limsup_{s\to 0} \frac{1}{s^2}\big\{d\big(x,\xi_v(-s)\big)^2 + d\big(x,\xi_v(s)\big)^2 - 2d(x,z)^2\big\} \le \frac{\partial^2}{\partial s^2}\big[L\big(\sigma(\cdot,0)\big)^2\big].$$

$\square$

Note that the analogue of the above theorem for $\mathcal{K} \ge k > 0$ (by replacing the right-hand side with $S^2\sqrt{k}r\cos(\sqrt{k}r)/\sin(\sqrt{k}r) + r\delta$) does not hold true. It is observed by considering $x = \eta(0)$ and $v = \dot{\eta}(l)$ for a geodesic η in a standard unit sphere (where $\mathcal{K} \equiv 1$, $\mathcal{T} \equiv 0$ and $\mathcal{S} \equiv 1$).

Remark 4.3 The three kinds of 'curvature bounds' in Theorem 4.2 appear for natural reasons. On one hand, the flag curvature measures how the space is curved like the sectional curvature in Riemannian geometry. On the other hand, the 2-uniform smoothness constant $\mathcal{S}$ controls the smoothness (concavity) of F on each tangent space (see Proposition 4.6 below), and the tangent curvature governs the variance of tangent spaces.

We put

$$h_0^s(\delta,S,r) := S^2 + r\delta, \quad h_k^s(\delta,S,r) := S^2\frac{\sqrt{k}r\cosh(\sqrt{k}r)}{\sinh(\sqrt{k}r)} + r\delta \ \text{ for } k>0. \tag{4.3}$$

Corollary 4.4 *Let (M,F), k, δ and S be as in Theorem 4.2 and take $r>0$. Then, for any $x \in M$, $z,z' \in B^+(x,r)\cap B^-(x,r)$, minimal geodesic $\eta : [0,1] \longrightarrow M$ from z to z' and for any $\tau \in [0,1]$, we have*

$$d\big(x,\eta(\tau)\big)^2 \ge (1-\tau)d(x,z)^2 + \tau d(x,z')^2 - (1-\tau)\tau h_k^s(\delta,S,3r)d(z,z')^2.$$

Proof. As $z,z' \in B^+(x,r)\cap B^-(x,r)$, we see

$$L(\eta) = d(z,z') \le d(z,x) + d(x,z') < 2r.$$

It implies $\eta \subset B^+(x,3r)$. Thus we deduce from Theorem 4.2 that, for any $\tau \in (0,1)$,

$$\limsup_{\varepsilon\to 0}\frac{1}{2\varepsilon^2}\big\{d\big(x,\eta(\tau-\varepsilon)\big)^2 + d\big(x,\eta(\tau+\varepsilon)\big)^2 - 2d\big(x,\eta(\tau)\big)^2\big\} \le h_k^s(\delta,S,3r)d(z,z')^2.$$

Therefore the function $\tau \longmapsto d(x,\eta(\tau))^2 - \tau^2 h_k^s(\delta,S,3r)d(z,z')^2$ is concave on [0,1], and hence we have

$$\begin{aligned}
&d\big(x,\eta(\tau)\big)^2 - \tau^2 h_k^s(\delta,S,3r)d(z,z')^2\\
&\ge (1-\tau)d(x,z)^2 + \tau\big\{d(x,z')^2 - h_k^s(\delta,S,3r)d(z,z')^2\big\}.
\end{aligned}$$

We obtain the required inequality by rearrangement. $\square$

Remark 4.5 The local version of Theorem 4.2 has been observed in [Sh3, Remark 15.1.4] (see also [Sh2, Theorem 5.1]). More precisely, he proves

$$\mathrm{Hess}\,\big[d(x,\cdot)\big](v) \le \frac{\mathbf{s}'_{-k}(r)}{\mathbf{s}_{-k}(r)}\big\{g_{T(r)}(v,v) - g_{T(r)}(v,T)^2\big\} - \mathcal{T}_{T(r)}(v)$$

for $z \in B^+(x, \mathrm{inj}(x))$. Then it follows from (3.3) that

$$\begin{aligned}
\mathrm{Hess}\left[d(x,\cdot)^2\right](v) &= 2r\,\mathrm{Hess}\left[d(x,\cdot)\right](v) + 2g_{T(r)}(v,T)^2 \\
&\le 2r\frac{\mathbf{s}'_{-k}(r)}{\mathbf{s}_{-k}(r)} g_{T(r)}(v,v) + 2\left(1 - r\frac{\mathbf{s}'_{-k}(r)}{\mathbf{s}_{-k}(r)}\right) g_{T(r)}(v,T)^2 + 2\delta r \\
&\le 2S^2 r\frac{\mathbf{s}'_{-k}(r)}{\mathbf{s}_{-k}(r)} + 2\delta r.
\end{aligned}$$

4.3 Minkowski spaces

Let us briefly comment on the case of Minkowski spaces. By a slight abuse of the notation, we denote a *Minkowski space* by $(\mathbb{R}^n, F)$, that is to say, $F : \mathbb{R}^n \longrightarrow [0,\infty)$ is a function satisfying the conditions (1–3) in Definition 2.1. By identifying each tangent space $T_v\mathbb{R}^n$ with the underlying space $\mathbb{R}^n$ in the canonical way, the function F also provides a Finsler structure of $\mathbb{R}^n$. Then $\mathcal{S}$ and $\mathcal{C}$ are constant because of the homogeneity of the space, and we denote these constants simply by $\mathcal{S}$ and $\mathcal{C}$, respectively.

Proposition 4.6 *Let $(\mathbb{R}^n, F)$ be a Minkowski space. Then the constant $\mathcal{S}$ coincides with the infimum of $S \ge 1$ satisfying*

$$F\left(\frac{v+w}{2}\right)^2 \ge \frac{1}{2}F(v)^2 + \frac{1}{2}F(w)^2 - \frac{S^2}{4}F(w-v)^2$$

for all $v, w \in \mathbb{R}^n$. Similarly, the constant $\mathcal{C}$ coincides with the infimum of $C \ge 1$ satisfying

$$F\left(\frac{v+w}{2}\right)^2 \le \frac{1}{2}F(v)^2 + \frac{1}{2}F(w)^2 - \frac{1}{4C^2}F(w-v)^2$$

for all $v, w \in \mathbb{R}^n$.

Proof. First of all, we remark that both the flag curvature $\mathcal{K}$ and the tangent curvature $\mathcal{T}$ are identically zero. Given $v \in \mathbb{R}^n$ with $F(v) = 1$ and $w \in \mathbb{R}^n \setminus 0$, we observe

$$\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon^2}\left\{F(w-\varepsilon v)^2 + F(w+\varepsilon v)^2 - 2F(w)^2\right\} = g_w(v,v)$$

by the definition of g_w. Hence we obtain

$$\begin{aligned}
&\sup_{v\in\mathbb{R}^n, F(v)=1}\ \sup_{w\in\mathbb{R}^n\setminus 0}\left[\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon^2}\left\{F(w-\varepsilon v)^2 + F(w+\varepsilon v)^2 - 2F(w)^2\right\}\right] \\
&= \sup_{v\in\mathbb{R}^n, F(v)=1}\ \sup_{w\in\mathbb{R}^n\setminus 0} g_w(v,v) = \mathcal{S}^2.
\end{aligned}$$

We also find

$$\begin{aligned}
&\inf_{v\in\mathbb{R}^n, F(v)=1}\ \inf_{w\in\mathbb{R}^n\setminus 0}\left[\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon^2}\left\{F(w-\varepsilon v)^2 + F(w+\varepsilon v)^2 - 2F(w)^2\right\}\right] \\
&= \left[\sup_{v\in\mathbb{R}^n, F(v)=1}\ \sup_{w\in\mathbb{R}^n\setminus 0} g_w(v,v)^{-1}\right]^{-1} = \mathcal{C}^{-2}.
\end{aligned}$$

These complete the proof through the discussion as in Corollary 4.4. $\square$

Therefore $\mathcal{S}$ and $\mathcal{C}$ coincide with the 2-uniform smoothness and convexity constants, respectively (see (1.1) and (1.2) in the introduction).

5 Uniform convexity

This section is devoted to the 2-uniform convexity as well as some observations related to Busemann's NPC-spaces.

5.1 Local uniform convexity

The following theorem is a counterpoint to Theorem 4.2. Although it has already been shown in [Sh3, Remark 15.1.4] (see also Remark 4.5), here we give a proof for completeness.

Theorem 5.1 (Uniform convexity) *Let (M,F) be a connected, forward geodesically complete C^∞-Finsler manifold. Assume that there are constants $k,\delta \ge 0$ and $C \ge 1$ such that*

$$\mathcal{K} \le k, \qquad \mathcal{T} \le \delta, \qquad \mathcal{C} \le C.$$

Take a point $x \in M$ and set $R := \mathrm{inj}(x)$ or $R := \min\{\mathrm{inj}(x), \pi/\sqrt{k}\}$ if $k > 0$. Then we have, for any $z \in B^+(x,R)$ and $v \in T_zM$ with $F(v) = 1$,

$$\begin{aligned}
&\liminf_{s\to 0} \frac{1}{2s^2}\Big\{d\big(x,\xi_v(-s)\big)^2 + d\big(x,\xi_v(s)\big)^2 - 2d(x,z)^2\Big\} \\
&\ge \begin{cases} C^{-2}\dfrac{\sqrt{k}r\cos(\sqrt{k}r)}{\sin(\sqrt{k}r)} - r\delta & \text{if } k > 0, \\ C^{-2} - r\delta & \text{if } k = 0, \end{cases}
\end{aligned}$$

where $\xi_v : (-\varepsilon,\varepsilon) \longrightarrow M$ is the geodesic with $\dot{\xi}_v(0) = v$ and $r := d(x,z)$.

Proof. We will omit some calculations common to the proof of Theorem 4.2. We can assume $z \neq x$ and $k > 0$. For small $\varepsilon > 0$, consider the variation $\sigma : [0,r] \times (-\varepsilon,\varepsilon) \longrightarrow M$ such that, for each $s \in (-\varepsilon,\varepsilon)$, the curve $t \longmapsto \sigma(t,s)$ is the unique minimal geodesic from x to $\xi_v(s)$. As $z \in B^+(x,\mathrm{inj}(x)) \setminus \{x\}$, the variation σ is C^∞ by Lemma 3.1. We put

$$\begin{aligned}
&\eta(t) := \sigma(t,0), \quad T(t,s) := \frac{\partial\sigma}{\partial t}(t,s), \quad U(t,s) := \frac{\partial\sigma}{\partial s}(t,s), \\
&T(t) := T(t,0) = \dot{\eta}(t), \quad U(t) := U(t,0).
\end{aligned}$$

Then $U(\cdot)$ is a Jacobi field along η, namely $D^{\dot{\eta}}_{\dot{\eta}} D^{\dot{\eta}}_{\dot{\eta}} U + R^{\dot{\eta}}(U,\dot{\eta})\dot{\eta} \equiv 0$. Thus it follows from the second variation formula (3.4) that

$$\begin{aligned}\frac{\partial^2}{\partial s^2}L\big(\sigma(\cdot,0)\big) &= \int_0^r \big\{g_{T(t)}(D_T^T U, D_T^T U) + g_{T(t)}(D_T^T D_T^T U, U)\big\}\,dt \\ &\quad + g_{T(r)}(D_U^T U, T) - \int_0^r \left\{\frac{\partial(F(T))}{\partial s}(t,0)\right\}^2 dt \\ &= g_{T(r)}(D_T^T U, U) - \mathcal{T}_{T(r)}(v) - \int_0^r \left\{\frac{1}{r}\frac{\partial}{\partial s}L\big(\sigma(\cdot,0)\big)\right\}^2 dt \\ &\geq g_{T(r)}(D_T^T U, U) - \delta - \frac{1}{r}\left\{\frac{\partial}{\partial s}L\big(\sigma(\cdot,0)\big)\right\}^2.\end{aligned}$$

Moreover, Theorem 3.2 shows that

$$g_{T(r)}(D_T^T U, U) \geq g_{T(r)}(v,v)\frac{\sqrt{k}\cos(\sqrt{k}r)}{\sin(\sqrt{k}r)} \geq C^{-2}\frac{\sqrt{k}\cos(\sqrt{k}r)}{\sin(\sqrt{k}r)}.$$

We consequently obtain

$$\begin{aligned}\frac{\partial^2}{\partial s^2}\Big[L\big(\sigma(\cdot,0)\big)^2\Big] &= 2r\frac{\partial^2}{\partial s^2}L\big(\sigma(\cdot,0)\big) + 2\left\{\frac{\partial}{\partial s}L\big(\sigma(\cdot,0)\big)\right\}^2 \\ &\geq 2r\left\{C^{-2}\frac{\sqrt{k}\cos(\sqrt{k}r)}{\sin(\sqrt{k}r)} - \delta\right\}.\end{aligned}$$

This completes the proof since $L(\sigma(\cdot,s)) = d(x,\xi_v(s))$. $\square$

We again remark that the analogue of the above theorem is false for $\mathcal{K} \leq -k < 0$ (see the paragraph following Theorem 4.2). If we put

$$h_0^c(\delta,C,r) := C^{-2} - r\delta, \quad h_k^c(\delta,C,r) := C^{-2}\frac{\sqrt{k}r\cos(\sqrt{k}r)}{\sin(\sqrt{k}r)} - r\delta \ \text{ for } k>0, \tag{5.1}$$

then the discussion as in Corollary 4.4 yields the following.

Corollary 5.2 *Let (M,F), k, δ, C, x and R be as in Theorem* 5.1 *and take $r \in (0,R/3)$. Then, for any $z,z' \in B^+(x,r) \cap B^-(x,r)$, minimal geodesic $\eta : [0,1] \longrightarrow M$ from z to z' and for any $\tau \in [0,1]$, we have*

$$d\big(x,\eta(\tau)\big)^2 \leq (1-\tau)d(x,z)^2 + \tau d(x,z')^2 - (1-\tau)\tau h_k^c(\delta,C,3r)d(z,z')^2.$$

We can take $R=\infty$ in Theorem 5.1 and Corollary 5.2 if (M,F) is simply connected and $k=0$. It is a consequence of the Finsler version of the Cartan-Hadamard theorem (cf. [BCS, Theorem 9.4.1]). Thus we further obtain the following.

Corollary 5.3 *Let (M,F) be a connected, simply connected, forward geodesically complete C^∞-Finsler manifold of Berwald type. Assume that $\mathcal{K} \leq 0$ and $\mathcal{C} \leq C$ for some constant $C \geq 1$. Then, for any $x,z,z' \in M$, minimal geodesic $\eta : [0,1] \longrightarrow M$ from z to z' and for any $\tau \in [0,1]$, we have*

$$d\big(x,\eta(\tau)\big)^2 \leq (1-\tau)d(x,z)^2 + \tau d(x,z')^2 - (1-\tau)\tau C^{-2}d(z,z')^2.$$

5.2 Local-to-global property in Busemann's NPC-spaces

We recall the definition of nonpositively curved spaces in the sense of Busemann (NPC-spaces for short, see [Bu] and [BH]). A (symmetric) metric space (X, d) is said to be *geodesic* if any two points $x, z \in X$ can be joined by a rectifiable curve $\eta : [0,1] \longrightarrow X$ whose arclength coincides with $d(x, z)$. A rectifiable curve $\eta : [0,1] \longrightarrow X$ is called a *geodesic* if it is locally minimizing and of constant speed. A geodesic η is said to be *minimal* if it is globally minimizing. Now, a geodesic space is said to be *NPC* (*in the sense of Busemann*) if, for any two minimal geodesics $\eta, \xi : [0,1] \longrightarrow X$ with $\eta(0) = \xi(0)$ and $\tau \in [0,1]$, we have

$$d\big(\eta(\tau), \xi(\tau)\big) \le \tau d\big(\eta(1), \xi(1)\big).$$

It is immediate by definition that any two points $x, z \in X$ in an NPC-space (X, d) admit a unique minimal geodesic between them and that (X, d) is contractible. A connected, simply connected, geodesically complete Riemannian manifold is NPC if and only if its sectional curvature is nonpositive everywhere. More generally, CAT(0)-spaces as well as strictly convex Banach spaces are NPC. However, the author does not know any reasonable condition for (positively or absolutely homogeneous) Finsler manifolds to be NPC.

Proposition 5.4 *Let (X, d) be an NPC-space in the sense of Busemann. Assume that there is a constant $C \ge 1$ such that, for any $x_0 \in X$, we find a positive number $r > 0$ for which we have*

$$d\big(x, \eta(\tau)\big)^2 \le (1-\tau)d(x,z)^2 + \tau d(x,z')^2 - (1-\tau)\tau C^{-2} d(z,z')^2 \tag{5.2}$$

for any $x, z, z' \in B(x_0, r)$, minimal geodesic $\eta : [0,1] \longrightarrow X$ from z to z' and for any $\tau \in [0,1]$. Then we have (5.2) globally, that is to say, (5.2) holds for any $x, z, z' \in X$, minimal geodesic $\eta : [0,1] \longrightarrow X$ from z to z' and for any $\tau \in [0,1]$.

Proof. Take arbitrary points $x, z, z' \in X$ and minimal geodesic $\eta : [0,1] \longrightarrow X$ from z to z'. Given $\tau \in (0,1)$, as (X, d) is NPC, we find a unique minimal geodesic $\xi : [0,1] \longrightarrow X$ from $\eta(\tau)$ to x. By assumption, for small $\varepsilon > 0$, it holds that

$$\begin{aligned} d\big(\xi(\varepsilon), \eta(\tau)\big)^2 \le (1-\tau) d\big(\xi(\varepsilon), \eta(\tau - \varepsilon\tau)\big)^2 + \tau d\Big(\xi(\varepsilon), \eta\big(\tau + \varepsilon(1-\tau)\big)\Big)^2 \\ - (1-\tau)\tau C^{-2} d\Big(\eta(\tau - \varepsilon\tau), \eta\big(\tau + \varepsilon(1-\tau)\big)\Big)^2. \end{aligned}$$

Combining this with the NPC-property, we observe

$$\begin{aligned} d\big(x, \eta(\tau)\big)^2 &= \varepsilon^{-2} d\big(\xi(\varepsilon), \eta(\tau)\big)^2 \\ &\le (1-\tau)d(x,z)^2 + \tau d(x,z')^2 - (1-\tau)\tau C^{-2} d(z,z')^2. \end{aligned}$$

□

We remark that the local-to-global theorem for the NPC-property is also known ([BH, Theorem II.4.1]).

Corollary 5.5 *Let (M,F) be a connected, absolutely homogeneous, geodesically complete C^∞-Finsler manifold. Assume that (M,F) is NPC in the sense of Busemann, and that there is a constant $C \ge 1$ such that $\mathcal{C} \le C$. Then, for any $x, z, z' \in M$, minimal geodesic $\eta : [0,1] \longrightarrow M$ from z to z' and for any $\tau \in [0,1]$, we have*

$$d\big(x, \eta(\tau)\big)^2 \le (1-\tau)d(x,z)^2 + \tau d(x,z')^2 - (1-\tau)\tau C^{-2} d(z,z')^2.$$

Proof. Apply Corollary 5.2 locally and observe that $\lim_{r\to 0} h_k^c(\delta, C, r) = C^{-2}$ for any $k \ge 0$. Then Proposition 5.4 completes the proof. □

5.3 L-convexity

We close the section with a remark on another type of convexity studied in [Oh2] (see also [Eg]). A geodesic space (X,d) is said to be *L-convex* for $L_1 \in [0,\infty)$ and $L_2 \in [0,\infty]$ if we have, for any $x \in X$, two minimal geodesics $\eta, \xi : [0,1] \longrightarrow X$ emanating from x and for any $\tau \in [0,1]$,

$$d\big(\eta(\tau), \xi(\tau)\big) \le \left(1 + L_1 \min\left\{\frac{d(x,\eta(1)) + d(x,\xi(1))}{2}, L_2\right\}\right) \tau d\big(\eta(1), \xi(1)\big).$$

What is the most important here is

$$\lim_{\eta(1),\xi(1)\to x} \left(1 + L_1 \min\left\{\frac{d(x,\eta(1)) + d(x,\xi(1))}{2}, L_2\right\}\right) = 1$$

(compare this with the properties (A) and (U) in [Ly]). NPC-spaces are obviously L-convex with $L_1 = 0$. Furthermore, a CAT(1)-space is L-convex if its diameter is less than π and if it does not contain any geodesic triangle of perimeter greater than 2π ([Oh2, Proposition 3.1(ii)]).

Let us take a connected, simply connected, forward geodesically complete C^∞-Finsler manifold (M,F) of nonpositive flag curvature. Then the Cartan-Hadamard theorem says that the exponential map $\exp_x : T_xM \longrightarrow M$ is a C^1-diffeomorphism for any $x \in M$. Given a point $x \in M$, a geodesic $\eta : [0,1] \longrightarrow M$ and $s \in [0,1]$, we denote by $\xi_s : [0,1] \longrightarrow M$ the unique geodesic from x to $\eta(s)$. If we put

$$T(t,s) := \frac{\partial \xi_s}{\partial t}(t), \qquad U(t,s) := \frac{\partial \xi_s}{\partial s}(t),$$

then each $U(\cdot, s)$ is a Jacobi field along ξ_s. We saw in the proof of Theorem 3.2 that

$$\frac{d^2}{dt^2}\left[g_{T(t,s)}\big(U(t,s), U(t,s)\big)^{1/2}\right] \ge 0. \tag{5.3}$$

However, this inequality does not imply the convexity of $d(\xi_0(\cdot), \xi_1(\cdot))$, for

$$g_{T(t,s)}\big(U(t,s), U(t,s)\big) \ne F\big(U(t,s)\big)^2.$$

What we can derive from (5.3) is, if we set $S := \sup_M \mathcal{S}$ and $C := \sup_M \mathcal{C}$,

$$\begin{aligned} d\big(\xi_0(\tau), \xi_1(\tau)\big) &\le C \int_0^1 g_{T(\tau,s)}\big(U(\tau,s), U(\tau,s)\big)^{1/2} \, ds \\ &\le C\tau \int_0^1 g_{T(1,s)}\big(U(1,s), U(1,s)\big)^{1/2} \, ds \le CS\tau d\big(\xi_0(1), \xi_1(1)\big). \end{aligned}$$

This 'quasi-convexity' has some applications in global theory (see [Eg]). However, from the viewpoints of metric geometry and of analysis on singular spaces, a somewhat stronger convexity (or concavity) is expected to hold in order to develop local theory (see, e.g., [Oh2] and [Ly]). It is desirable if we can establish the L-convexity of (M, F) with appropriate bounds on L_1 (and L_2) in terms of the curvatures and of the 2-uniform convexity and smoothness constants.

6 Almost everywhere second order differentiability of semi-convex functions

This section is devoted to the almost everywhere existence of the second order differentials of semi-convex functions on Finsler manifolds. Such a differentiability has been established by Alexandrov [Al] and Bangert [Ba] in the Euclidean and Riemannian cases, respectively. Bangert in fact treats a more general situation, and we will see that his theorem also covers our Finsler setting.

6.1 Bangert's theorem

Let us take an n-dimensional C^∞-manifold M and define a class $\mathfrak{F}(M)$ of functions on M as follows. A function $f : M \longrightarrow \mathbb{R}$ is an element of $\mathfrak{F}(M)$ if, for any point $x \in M$, there are a local coordinate system $\Phi : U \longrightarrow \mathbb{R}^n$ and a C^∞-function $h : U \longrightarrow \mathbb{R}$ on an open set $U \subset M$ containing x such that the function $(f + h) \circ \Phi^{-1} : \Phi(U) \longrightarrow \mathbb{R}$ is convex in the usual Euclidean sense. Before stating Bangert's theorem, we recall two notions on the differentials of functions.

Definition 6.1 (Subdifferentials) Let $f \in \mathfrak{F}(M)$, $x \in M$ and let $\Phi : U \longrightarrow \mathbb{R}^n$ be a local coordinate system on an open set $U \subset M$ with $x \in U$ and $\Phi(x) = 0$. Then a co-vector $\alpha \in T_x^*M$ is called a *subgradient* of f at x if we have

$$f\big(\Phi^{-1}(u)\big) \ge f(x) + \alpha\big([d(\Phi^{-1})_0](u)\big) + o(\|u\|)$$

for $u \in \mathbb{R}^n$. Here we regard u as an element of $T_0\mathbb{R}^n$ in the second term in the right-hand side, and denote by $\|\cdot\|$ the Euclidean norm of $\mathbb{R}^n$. The set of all subgradients at x is called the *subdifferential* of f at x and denoted by $\partial_* f(x) \subset T_x^*M$.

Note that the definition of the subgradient does not depend on the choice of a local coordinate system Φ. It is known that a function $f \in \mathfrak{F}(M)$ admits a (not necessarily unique) subgradient everywhere. In Euclidean spaces and Riemannian manifolds, the dual of α in Definition 6.1 is usually called the subgradient, and the corresponding subdifferential $\partial f(x)$ is a subset of T_xM. Also in Finsler manifolds, it is possible to define the subgradient as an element of T_xM through the Legendre transform $\mathfrak{L}_x : T_x^*M \longrightarrow T_xM$. It actually coincides with the gradient vector of f if f is differentiable at x, namely we have $\partial f(x) = \{\operatorname{grad} f(x)\}$.

Definition 6.2 (Second order differentials) Let $f \in \mathfrak{F}(M)$, $x \in M$ and let $\Phi : U \longrightarrow \mathbb{R}^n$ be a local coordinate system on an open set $U \subset M$ with $x \in U$ and $\Phi(x) = 0$. Then f is said to be *second order differentiable* at x if there is a linear map $H : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that

$$\sup_{\alpha(x)\in\partial_* f(x),\alpha(z)\in\partial_* f(z)} \left\| \left[d(\Phi^{-1})_u\right]^* \big(\alpha(z)\big) - \left[d(\Phi^{-1})_0\right]^* \big(\alpha(x)\big) - Hu \right\| = o(\|u\|)$$

for $u = \Phi(z) \in \mathbb{R}^n$, where we identify $T_u^*\mathbb{R}^n$, $T_0^*\mathbb{R}^n$ and $\mathbb{R}^n$ in the left-hand side.

Observe that the above definition implies that the subdifferential $\partial_* f(x)$ at x consists of a single element, and hence f is differentiable at x.

Theorem 6.3 ([Ba, Satz (4.4)]) *Let M be a C^∞-manifold. Then every function $f \in \mathfrak{F}(M)$ is second order differentiable almost everywhere.*

Here 'almost everywhere' means that the set of second order differentiable points has a full n-dimensional Lebesgue measure in $\Phi(U)$ for any local coordinate system $\Phi : U \longrightarrow \mathbb{R}^n$ on $U \subset M$.

Remark 6.4 If M is equipped with a C^∞-Riemannian metric g, then we can write the second order Taylor expansion by using the exponential map as

$$f(\exp_x v) = f(x) + g\big(\operatorname{grad} f(x), v\big) + \frac{1}{2} \operatorname{Hess} f(v, v) + o(\|v\|_g^2),$$

where $v \in T_xM$ and $\|v\|_g^2 := g(v, v)$. However, it is not the case of Finsler structures because the exponential map is only C^1 at the zero section for general Finsler manifolds.

A function $f : M \longrightarrow \mathbb{R}$ on a Finsler (or Riemannian) manifold (M, F) is said to be *convex* if it is convex along any geodesic on M, namely

$$f\big(\eta(t)\big) \le (1-t)f\big(\eta(0)\big) + tf\big(\eta(1)\big)$$

holds for any geodesic $\eta : [0, 1] \longrightarrow M$ and $t \in [0, 1]$. Similarly, a function $f : M \longrightarrow \mathbb{R}$ is said to be λ-*convex* for $\lambda \in \mathbb{R}$ if we have

$$f\big(\eta(t)\big) \le (1-t)f\big(\eta(0)\big) + tf\big(\eta(1)\big) - \frac{\lambda}{2}(1-t)td\big(\eta(0), \eta(1)\big)^2$$

for any geodesic $\eta : [0, 1] \longrightarrow M$ and $t \in [0, 1]$. We say that f is *semi-convex* if, for any $x \in M$, there are an open neighborhood U of x and a constant $\lambda = \lambda(x) \in \mathbb{R}$ for which $f|_U$ is λ-convex.

Theorem 6.5 ([Ba, Satz (2.3)]) *Let (M, F) be a C^∞-Riemannian manifold. Then every function $f : M \longrightarrow \mathbb{R}$ which is convex with respect to g is an element of $\mathfrak{F}(M)$.*

Bangert's proof effectively uses Greene and Wu's approximation theory ([GW1] and [GW2]), and we will manage to extend it to the Finsler case. Theorem 6.5 immediately implies that semi-convex functions on a Riemannian manifold M are also in $\mathfrak{F}(M)$, for the squared distance function $d(x, \cdot)^2$ from a point x is C^∞ and 1-convex on a neighborhood of x (consider $f - \min\{0, \lambda(x)\}d(x, \cdot)^2$). In the remainder of the section, we prove the following:

Theorem 6.6 *Let (M, F) be a C^∞-Finsler manifold. Then every function $f : M \longrightarrow \mathbb{R}$ which is semi-convex with respect to F is an element of $\mathfrak{F}(M)$. In particular, f is second order differentiable almost everywhere.*

6.2 Approximations of convex functions

Let $f : M \longrightarrow \mathbb{R}$ be a convex function on a C^∞-Finsler manifold (M, F). According to Greene and Wu's technique, we will construct a family of nearly convex C^∞-functions which approximates f.

Fix a point $x \in M$ and $r > 0$ such that the forward closed ball $\bar{B}^+(x, r)$ is compact. Denote by L the *Lipschitz constant* of $f|_{B^+(x,r)}$ in the sense that

$$f(y) - Ld(z, y) \le f(z) \le f(y) + Ld(y, z)$$

for all $y, z \in B^+(x, r)$. We choose a C^∞-Riemannian metric g on $B^+(x, r)$ which is bi-Lipschitz equivalent to F. One way to construct such a metric is putting $g := g_V$ for some non-vanishing C^∞-vector field V on $B^+(x, r)$. Then we have $\mathcal{C}(z)^{-2} \le g_V(v, v)/F(v)^2 \le \mathcal{S}(z)^2$ for all $v \in T_zM \setminus 0$ with $z \in B^+(x, r)$ (recall (4.1) and (4.2)). It might be also possible to find g with a universal bound on $\sup_{v \in \pi^{-1}(B^+(x,r)) \setminus 0} \{g(v, v)/F(v)^2, F(v)^2/g(v, v)\}$ in terms only of the dimension of M like John's theorem ([Jo]).

Fix a nonnegative C^∞-function $\phi : \mathbb{R} \longrightarrow [0, \infty)$ satisfying $\operatorname{supp} \phi \subset [1/2, 1]$ and $\int_{\mathbb{R}^n} \phi(\|v\|) \, dv = 1$, where dv stands for the n-dimensional Lebesgue measure and $\|v\|$ denotes the Euclidean norm of v. For large $i \in \mathbb{N}$ and $z \in B^+(x, r)$, we define

$$f_i(z) := i^n \int_{T_zM} f(\exp_z^g v) \phi\big(i\|v\|_g\big) \, d\mathsf{L}_g(v),$$

where $\exp^g$ is the exponential map with respect to g and L_g is the Lebesgue measure on T_zM induced from g. It is immediate that f_i is C^∞ and converges to f uniformly on $B^+(x, r)$ as i goes to infinity.

Given an F-unit vector $v \in \pi^{-1}(B^+(x, r))$ (i.e., $F(v) = 1$), let $\eta : [-\tau, \tau] \longrightarrow M$ be the geodesic with respect to F with $\dot{\eta}(0) = v$, where $\tau > 0$ is chosen small relative to r and to the infimum on $B^+(x, r)$ of the injectivity radius with respect to F (it is possible because the function $z \longmapsto \operatorname{inj}(z)$ is continuous). Put $B(\delta) := \{w \in T_{\eta(-\tau)}M \mid \|w\|_g \le \delta\}$ for small $\delta > 0$, and define the map $Q_1 : B(\delta) \longrightarrow C^\infty([-\tau, \tau], M)$ as

$$Q_1(w)(t) := \exp_{\eta(t)}^g \big(\Xi_{-\tau,t}^g(w)\big),$$

where $\Xi^g_{-\tau,t} : T_{\eta(-\tau)}M \longrightarrow T_{\eta(t)}M$ is the parallel translation along η with respect to g. Note that

$$f_i\big(\eta(t)\big) = i^n \int_{T_{\eta(-\tau)}M} f\big(Q_1(w)(t)\big)\phi\big(i\|w\|_g\big)\, d\mathsf{L}_g(w).$$

We consider the C^∞-topology on $C^\infty([-\tau,\tau],M)$ metrized by using g, and observe that Q_1 is continuous and $Q_1(0)=\eta$. (Here is a reason why we used g in the definition of f_i instead of F. Recall that $\exp^F$ is only C^1 at the zero section.) Hence the Lipschitz constant of f_i with respect to F converges to L as i diverges to infinity (see [GW2, Lemma 8]).

Since $F(\dot\eta(-\tau))=1$, by taking δ small enough, it holds that $F([\partial Q_1(w)/\partial t](-\tau)) \in [1/2,3/2]$ for all $w\in B(\delta)$. Therefore the map $Q_2 : B(\delta) \longrightarrow C^{-\infty}([-\tau,\tau],M)$ defined by

$$Q_2(w)(t) = \exp^F_{Q_1(w)(-\tau)}\left((t+\tau)\frac{\partial Q_1(w)}{\partial t}(-\tau)\right)$$

is continuous (see Lemma 3.1). Since $Q_1(0)=Q_2(0)=\eta$ as well as $[\partial Q_1(w)/\partial t](-\tau) = [\partial Q_2(w)/\partial t](-\tau)$, given $\varepsilon>0$, there is $\delta>0$ such that $d_F(Q_1(w)(t),Q_2(w)(t)) \le \varepsilon\tau^2$ holds for all $w\in B(\delta)$ and $t\in[-\tau,\tau]$. Moreover, δ can be chosen uniformly in F-unit vectors $v\in\pi^{-1}(B^+(x,r))$ since $\bar B^+(x,r)$ is compact. Let

$$m : \exp^g_{\eta(-\tau)}\big(B(\delta)\big) \times \exp^g_{\eta(\tau)}\Big(\Xi^g_{-\tau,\tau}\big(B(\delta)\big)\Big) \longrightarrow M$$

be the map such that $m(z,z')=\xi(1/2)$, where $\xi : [0,1] \longrightarrow M$ is the unique minimal geodesic from z to z' with respect to F. Then m is C^∞ (by the same reason as the continuity of Q_2) and $m(Q_2(w)(-\tau),Q_2(w)(\tau)) = Q_2(w)(0)$, and hence we have, by taking smaller δ if necessary,

$$d_F\Big(m\big(Q_1(w)(-\tau),Q_1(w)(\tau)\big),Q_2(w)(0)\Big) \le \varepsilon\tau^2$$

for all F-unit vectors $v\in\pi^{-1}(B^+(x,r))$ and all $w\in B(\delta)$ (see [GW1, Proposition]).

Thus we obtain, for any F-unit vector $v\in\pi^{-1}(B^+(x,r))$ and $i\ge\delta^{-1}$,

$$\begin{aligned}
&f_i\big(\eta(-\tau)\big) + f_i\big(\eta(\tau)\big) - 2f_i\big(\eta(0)\big)\\
&= i^n \int_{T_{\eta(-\tau)}M} \Big\{f\big(Q_1(w)(-\tau)\big) + f\big(Q_1(w)(\tau)\big) - 2f\big(Q_1(w)(0)\big)\Big\}\phi\big(i\|w\|_g\big)\, d\mathsf{L}_g(w)\\
&= i^n \int_{T_{\eta(-\tau)}M} \bigg[\Big\{f\big(Q_1(w)(-\tau)\big) + f\big(Q_1(w)(\tau)\big) - 2f\Big(m\big(Q_1(w)(-\tau),Q_1(w)(\tau)\big)\Big)\Big\}\\
&\qquad + 2\Big\{f\Big(m\big(Q_1(w)(-\tau),Q_1(w)(\tau)\big)\Big) - f\big(Q_2(w)(0)\big)\Big\}\\
&\qquad + 2\Big\{f\big(Q_2(w)(0)\big) - f\big(Q_1(w)(0)\big)\Big\}\bigg]\,\phi\big(i\|w\|_g\big)\, d\mathsf{L}_g(w)\\
&\ge 0 - 4L\varepsilon\tau^2\cdot i^n \int_{T_{\eta(-\tau)}M} \phi\big(i\|w\|_g\big)\, d\mathsf{L}_g(w) = -4L\varepsilon\tau^2.
\end{aligned}$$

Here the third inequality follows from the convexity of f with respect to F. Therefore we have $\mathrm{Hess}^F f_i(v) \ge -4L\varepsilon$ for any F-unit vector $v\in\pi^{-1}(B^+(x,r))$. We consequently deduce the following analogue of [GW1, Theorem 2] (see also [GW2, Lemma 8]).

Proposition 6.7 *Let $f : M \longrightarrow \mathbb{R}$ be a convex function on a C^∞-Finsler manifold (M,F). Then, for any $x \in M$ and $r > 0$ such that $\bar{B}^+(x,r)$ is compact, there exists a family $\{f_i\}_{i\in\mathbb{N}}$ of C^∞-functions on $B^+(x,r)$ satisfying the following conditions*:

(i) *The function f_i uniformly converges to $f|_{B^+(x,r)}$ as i diverges to infinity.*

(ii) *The Lipschitz constant of f_i tends to that of $f|_{B^+(x,r)}$ as i goes to infinity.*

(iii) *It holds that*

$$\liminf_{i\to\infty} \left(\inf_{v\in\pi^{-1}(B^+(x,r)),F(v)=1} \mathrm{Hess}^F f_i(v) \right) \ge 0.$$

6.3 Proof of Theorem 6.6

We first consider a convex function $f : M \longrightarrow \mathbb{R}$ and use the approximation $\{f_i\}_{i\in\mathbb{N}}$ as in Proposition 6.7 to proceed along Bangert's argument. We may assume $F(\mathrm{grad}^F f_i(z)) \le L + i^{-1}$ and $\mathrm{Hess}^F f_i(v) \ge -i^{-1}F(v)^2$ for all $z \in B^+(x,r)$ and $v \in \pi^{-1}(B^+(x,r)) \setminus 0$. Recall that L is the Lipschitz constant of $f|_{B^+(x,r)}$. Moreover, we can choose smaller $r > 0$ satisfying $r \ll \inf_{z\in B^+(x,r)} \mathrm{inj}(z)$ because the function $z \longmapsto \mathrm{inj}(z)$ is continuous. Fix a local coordinate system $\Phi : U \longrightarrow \mathbb{R}^n$ on an open set $U \supset \bar{B}^+(x,r)$ and denote by g the C^∞-Riemannian metric on U induced through Φ from the standard Euclidean structure of $\mathbb{R}^n$.

Given a g-unit vector $v \in \pi^{-1}(B^+(x,r/2))$, let $\xi_v : [-\tau,\tau] \longrightarrow M$ be the geodesic with respect to F such that $\dot{\xi}_v(0) = v$ and $d_F(x,\xi_v(-\tau)) = r$. Then we define $V := \mathrm{grad}^F[d_F(\xi_v(-\tau),\cdot)]$ and observe that V is a C^∞-vector field on $B^+(x,r/2)$. Thus we can define $g_v := g_V$ as a C^∞-Riemannian metric of $B^+(x,r/2)$.

Lemma 6.8 *The curve ξ_v is a geodesic with respect to g_v.*

Proof. This lemma is a consequence of [Sh3, Lemma 6.2.1, (6.14)]. Here we give a direct proof for thoroughness. In this proof, we denote quantities with respect to g_v by $\hat{g}_{ij}$, $\hat{\gamma}^i{}_{jk}$ and so forth.

We first observe that $\hat{g}_{ij}(\pi(V)) = g_{ij}(V)$ and

$$\begin{aligned}\frac{\partial \hat{g}_{ij}}{\partial x^k}\big(\pi(V)\big) &= \frac{\partial g_{ij}}{\partial x^k}(V) + \sum_m \frac{\partial g_{ij}}{\partial y^m}(V)\frac{\partial V^m}{\partial x^k}\big(\pi(V)\big) \\ &= \frac{\partial g_{ij}}{\partial x^k}(V) + \frac{2}{F}\sum_m A_{ijm}(V)\frac{\partial V^m}{\partial x^k}\big(\pi(V)\big).\end{aligned}$$

Combining this with $\sum_i A_{ijk}(V)V^i(\pi(V)) = 0$ (cf. [BCS, (1.4.6)]), we see

$$\sum_{j,k} \hat{\gamma}^i{}_{jk}\big(\pi(V)\big)V^jV^k = \sum_{j,k} \gamma^i{}_{jk}(V)V^jV^k = \sum_{j,k} \Gamma^i{}_{jk}(V)V^jV^k.$$

Therefore the geodesic equation $D_V^V V = 0$ is common to F and g_v. $\square$

As $\bar{B}^+(x,r)$ is compact, the following quantities are finite:

$$R_1 := \sup_{v \in \pi^{-1}(B^+(x,r/2)) \setminus 0} \frac{F(v)^2}{g(v,v)}, \qquad R_2 := \sup_{v \in \pi^{-1}(B^+(x,r/2)), g(v,v)=1} \left\| \nabla_v^{g_v} \dot{c}_v(0) \right\|_{g_v},$$

where we set $c_v(t) := \exp^g tv$ and denote by ∇^{g_v} the Levi-Civita connection with respect to g_v. For any g-unit vector $v \in T_zM$ with $z \in B^+(x,r/2)$, we observe

$$\begin{aligned}
\operatorname{Hess}^g f_i(v,v) &= \frac{d^2(f_i \circ c_v)}{dt^2}(0) = \frac{d}{dt}\Big|_{t=0} \Big[g_v\big(\dot{c}_v(t), \operatorname{grad}^{g_v} f_i \circ c_v(t)\big) \Big] \\
&= g_v\big(\nabla_v^{g_v} \dot{c}_v(0), \operatorname{grad}^{g_v} f_i(z)\big) + g_v\big(v, (\nabla_v^{g_v} \operatorname{grad}^{g_v} f_i)(z)\big).
\end{aligned}$$

Note that

$$\begin{aligned}
g_v\big(\nabla_v^{g_v} \dot{c}_v(0), \operatorname{grad}^{g_v} f_i(z)\big) &\ge -\left\| \nabla_v^{g_v} \dot{c}_v(0) \right\|_{g_v} \cdot \left\| \operatorname{grad}^{g_v} f_i(z) \right\|_{g_v} \\
&\ge -R_2(L + i^{-1}), \\
g_v\big(v, (\nabla_v^{g_v} \operatorname{grad}^{g_v} f_i)(z)\big) &= \operatorname{Hess}^{g_v} f_i(v,v) = \frac{d^2(f_i \circ \xi_v)}{dt^2}(0) = \operatorname{Hess}^F f_i(v) \\
&\ge -i^{-1} F(v)^2 \ge -i^{-1} R_1.
\end{aligned}$$

Therefore we find

$$\operatorname{Hess}^g f_i(v,v) \ge -R_2(L + i^{-1}) - R_1 i^{-1}.$$

Now we define the function $h(z) := \{R_2(L+1) + R_1\} d_g(x,z)^2$ for $z \in B^+(x,r/2)$. Since $\operatorname{Hess}^g[d_g(x,\cdot)^2](v,v) = 2$, we obtain $\operatorname{Hess}^g(f_i + h)(v,v) \ge 0$ for any g-unit vector $v \in \pi^{-1}(B^+(x,r/2))$, and hence $f_i + h$ is convex with respect to g. It follows from Proposition 6.7(i) that $f + h$ is also convex with respect to g, and it implies $f \in \mathfrak{F}(M)$. Thus we complete the proof of Theorem 6.6 for convex functions.

For a semi-convex function $f : M \longrightarrow \mathbb{R}$ and a point $x \in M$, we take small $r > 0$ such that f is λ-convex on $B^+(x,r)$ for some $\lambda \le 0$ and that

$$r \ll \inf_{z \in B^+(x,r)} \operatorname{inj}(z), \qquad h_k^c(\delta, C, 3r) \ge C^{-2}/2,$$

where $k, \delta \ge 0$ and $C \ge 1$ are chosen so as to satisfy $\mathcal{K} \le k$, $\mathcal{T} \le \delta$ and $\mathcal{C} \le C$ on $B^+(x,3r)$. Recall (5.1) for the definition of h_k^c. We set $B^\pm(x,r) := B^+(x,r) \cap B^-(x,r)$ for a while for brevity. Then the squared distance function $h := d_F(z,\cdot)^2$ from a point $z \in B^\pm(x,r) \setminus \bar{B}^\pm(x,r/2)$ is C^∞ on $B^\pm(x,r/2)$ and satisfies $\operatorname{Hess}^F h(v) \ge C^{-2}$ by Theorem 5.1 . Note indeed that

$$d(x,z) + d(z,z') < r + d(z,x) + d(x,z') < 5r/2$$

for any $z' \in B^\pm(x,r/2)$, and hence the unique minimal geodesic from z to z' is contained in $B^+(x,3r)$. Therefore $f - \lambda C^2 h$ is convex on $B^\pm(x,r/2)$, and thus $f - \lambda C^2 h \in \mathfrak{F}(B^\pm(x,r/2))$. As h is smooth, it shows $f \in \mathfrak{F}(M)$ and completes the proof of Theorem 6.6.

7 Almost everywhere second order differentiability of c-concave functions

In this final section, we verify that a c-concave function φ on a Finsler manifold with the quadratic cost function $c(x,y) = d(x,y)^2/2$ is second order differentiable almost everywhere. To do so, we will follow a Riemannian discussion in [CMS]. More precisely, we use Theorem 4.2 to show that $-\varphi$ is semi-convex in the sense of Section 6, and then Theorem 6.6 is applied.

Throughout the section, let (M,F) be a connected, forward geodesically complete C^∞-Finsler manifold, and fix the *quadratic cost function* $c(x,y) := d(x,y)^2/2$. This is the most fundamental cost function, however, c is not necessarily symmetric, for (M,F) is only positively homogeneous. This is one of the most natural situations where we encounter a nonsymmetric cost function.

Let $X, Y \subset M$ be two compact sets. Then, given an arbitrary function $\varphi : X \longrightarrow \mathbb{R} \cup \{-\infty\}$, we define its *c-transform* $\varphi^c : Y \longrightarrow \mathbb{R} \cup \{-\infty\}$ relative to (X,Y) by

$$\varphi^c(y) := \inf_{x \in X} \{c(x,y) - \varphi(x)\}.$$

Similarly, we define the c-transform of a function $\psi : Y \longrightarrow \mathbb{R} \cup \{-\infty\}$ relative to (X,Y) by $\psi^c(x) := \inf_{y \in Y}\{c(x,y) - \psi(y)\}$ for $x \in X$. (Be careful of the order of x and y in c.)

Definition 7.1 (c-concave functions) Let $X, Y \subset M$ be two compact sets. Then a function $\varphi : X \longrightarrow \mathbb{R} \cup \{-\infty\}$ which is not identically $-\infty$ is said to be *c-concave* relative to (X,Y) if there is a function $\psi : Y \longrightarrow \mathbb{R} \cup \{-\infty\}$ whose c-transform ψ^c relative to (X,Y) coincides with φ.

The class of c-concave functions is quite an important object in mass transport theory. Roughly speaking, any optimal transport between two probability measures can be described as the transport along the gradient vector field of some c-concave function. Therefore the almost everywhere second order differentiability of c-concave functions plays a crucial role in investigating the behavior of such an optimal transport, and is actually one of the key technical ingredients in Cordero-Erausquin, McCann and Schmuckenschläger's work on Riemannian interpolation inequalities (see [CMS, Proposition 3.14]). See also [Oh6] for a recent generalization to Finsler manifolds. We refer to [Br], [CMS], [Mc], [RR], [Vi1], [Vi2] and the references therein for more details and further reading.

We summerize some basic properties of c-concave functions in the next lemma, and prove it for completeness.

Lemma 7.2 *Take two compact sets $X, Y \subset M$ and a function $\varphi : X \longrightarrow \mathbb{R} \cup \{-\infty\}$. Then the following hold*:

(i) *We have $\varphi \le \varphi^{cc}$ and $\varphi^c = \varphi^{ccc}$.*

(ii) *Assume that φ is not identically $-\infty$. Then φ is c-concave if and only if $\varphi = \varphi^{cc}$.*

(iii) *If φ is c-concave, then it is Lipschitz continuous.*

Proof. (i) For any $x \in X$ and $y \in Y$, it follows from the definition of φ^c that $\varphi(x) \le c(x,y) - \varphi^c(y)$. By taking the infimum over $y \in Y$, we find $\varphi(x) \le \varphi^{cc}(x)$. In particular, we see $\varphi^c \le \varphi^{ccc}$. Moreover, given $y \in Y$, we observe

$$\begin{aligned}\varphi^{ccc}(y) &= \inf_{x\in X}\sup_{z\in Y}\inf_{w\in X}\big\{c(x,y) - c(x,z) + c(w,z) - \varphi(w)\big\}\\ &\le \inf_{x\in X}\big\{c(x,y) - \varphi(x)\big\} = \varphi^c(y),\end{aligned}$$

where the second inequality is deduced by choosing $w = x$. Thus we obtain $\varphi^c(y) = \varphi^{ccc}(y)$.

(ii) If φ is c-concave, then $\varphi = \psi^c$ for some function $\psi : Y \longrightarrow \mathbb{R} \cup \{-\infty\}$. Hence we have $\varphi = \psi^c = \psi^{ccc} = \varphi^{cc}$ by (i). The converse is clear by the definition of c-concave functions.

(iii) The function c is bounded on $X \times Y$ since X and Y are compact. Thus φ^c is bounded from above (otherwise, $\varphi = \varphi^{cc}$ is identically $-\infty$), and hence φ takes finite value everywhere. For $x \in X$ and $\varepsilon > 0$, take a point $y \in Y$ where $\varphi(x) \ge c(x,y) - \varphi^c(y) - \varepsilon$ holds. By the definition of $\varphi^{cc} = \varphi$, $\varphi(w) \le c(w,y) - \varphi^c(y)$ holds for any $w \in X$, and hence

$$\begin{aligned}\varphi(w) - \varphi(x) &\le c(w,y) - c(x,y) + \varepsilon\\ &= \frac{1}{2}\big\{d(w,y) + d(x,y)\big\}\big\{d(w,y) - d(x,y)\big\} + \varepsilon\\ &\le \Big\{\sup_{z\in X, y\in Y} d(z,y)\Big\} \cdot d(w,x) + \varepsilon.\end{aligned}$$

As ε is arbitrary, this derives $\varphi(w) - \varphi(x) \le C d(w,x)$ with $C = \sup_{z\in X, y\in Y} d(z,y)$ which could be adopted as the definition of the C-Lipschitz continuity of the function $-\varphi$ in our nonsymmetric metric space. Moreover, we obtain

$$\begin{aligned}|\varphi(w) - \varphi(x)| &\le \sup_{z\in X, y\in Y} d(z,y) \cdot \max\big\{d(w,x), d(x,w)\big\}\\ &\le \sup_{z\in X, y\in Y} d(z,y) \cdot \sup_{z\in \bar{B}^+(w, d(w,x))} \mathcal{C}(z) \cdot d(w,x).\end{aligned}$$

Here the second inequality follows from the observation that, for a minimal geodesic $\eta : [0,1] \longrightarrow M$ from w to x,

$$\begin{aligned}d(x,w) &\le \int_0^1 F\big(-\dot{\eta}(t)\big)\,dt \le \int_0^1 \mathcal{C}\big(\eta(t)\big) g_{\dot{\eta}(t)}\big(-\dot{\eta}(t), -\dot{\eta}(t)\big)^{1/2}\,dt\\ &= \int_0^1 \mathcal{C}\big(\eta(t)\big) F\big(\dot{\eta}(t)\big)\,dt \le \sup_{t\in[0,1]} \mathcal{C}\big(\eta(t)\big) \cdot d(w,x).\end{aligned}$$

The terms $\sup_{z\in X, y\in Y} d(z,y)$ and $\sup_{w,x\in X, z\in \bar{B}^+(w,d(w,x))} \mathcal{C}(z)$ are finite since X and Y are compact. $\square$

By virtue of Lemma 7.2(iii), we can restrict φ and ψ $(= \varphi^c)$ in Definition 7.1 to Lipschitz continuous, real-valued functions without loss of generality. In particular, the infimum $\varphi(x) = \varphi^{cc}(x) = \inf_{y\in Y}\{c(x,y) - \varphi^c(y)\}$ is attained at some point $y \in Y$.

Definition 7.3 (c-superdifferentials) Let $X, Y \subset M$ be compact sets and $\varphi : X \longrightarrow \mathbb{R}$ be a c-concave function relative to (X, Y). Then the *c-superdifferential* of φ at a point $x \in X$ is the nonempty set

$$\begin{aligned} \partial^c \varphi(x) &:= \{y \in Y \mid \varphi(x) = c(x,y) - \varphi^c(y)\} \\ &= \{y \in Y \mid c(x,y) - \varphi(x) \le c(z,y) - \varphi(z) \text{ for any } z \in X\}. \end{aligned}$$

Theorem 7.4 *Let (M, F) be a connected, forward geodesically complete C^∞-Finsler manifold and $c(x,y) = d(x,y)^2/2$ be the quadratic cost function. Take a compact set $Y \subset M$ and an open set $U \subset M$ whose closure $\overline{U}$ is compact. Then, for any c-concave function $\varphi : \overline{U} \longrightarrow \mathbb{R}$ relative to $(\overline{U}, Y)$, the function $-\varphi$ is semi-convex on U with respect to F. In particular, φ is second order differentiable almost everywhere on U.*

Proof. Fix a minimal geodesic $\eta : [0,1] \longrightarrow U$ and take $y \in \partial^c\varphi(\eta(1/2)) \subset Y$. On one hand, by the definition of the c-superdifferential $\partial^c\varphi(\eta(1/2))$, we observe

$$\begin{aligned} \varphi\big(\eta(0)\big) &\le \varphi\big(\eta(1/2)\big) + \frac{1}{2} d\big(\eta(0), y\big)^2 - \frac{1}{2} d\big(\eta(1/2), y\big)^2, \\ \varphi\big(\eta(1)\big) &\le \varphi\big(\eta(1/2)\big) + \frac{1}{2} d\big(\eta(1), y\big)^2 - \frac{1}{2} d\big(\eta(1/2), y\big)^2. \end{aligned}$$

On the other hand, it follows from Theorem 4.2 (see also Corollary 4.4) that

$$d\big(y, \eta(1/2)\big)^2 \ge \frac{1}{2} d\big(y, \eta(0)\big)^2 + \frac{1}{2} d\big(y, \eta(1)\big)^2 - \frac{\Lambda}{4} d\big(\eta(0), \eta(1)\big)^2,$$

where we set $\Lambda := h_k^s(\delta, S, r)$ for $r := \sup_{w \in Y, x \in U} d(w, x)$, $k, \delta \ge 0$ and $S \ge 1$ satisfying $\mathcal{K} \ge -k$, $\mathcal{T} \ge -\delta$ and $\mathcal{S} \le S$ on $\bigcup_{w \in Y} B^+(w, r)$. Recall (4.3) for the definition of h_k^s and note that Λ is finite since $\overline{U}$ and Y are compact. Therefore we obtain

$$\begin{aligned} -\varphi\big(\eta(1/2)\big) &\le -\frac{\varphi(\eta(0)) + \varphi(\eta(1))}{2} + \frac{d(\eta(0), y)^2 + d(\eta(1), y)^2}{4} - \frac{1}{2} d\big(\eta(1/2), y\big)^2 \\ &\le -\frac{1}{2}\varphi\big(\eta(0)\big) - \frac{1}{2}\varphi\big(\eta(1)\big) + \frac{\Lambda}{8} d\big(\eta(0), \eta(1)\big)^2. \end{aligned}$$

As Λ depends only on U and Y, $-\varphi$ is $(-\Lambda)$-convex on U. By combining this with Theorem 6.6, we complete the proof. $\square$

References

[Al] A. D. Alexandrov, *Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it* (Russian), Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser. **6** (1939), 3–35.

[BCL] K. Ball, E. A. Carlen and E. H. Lieb, *Sharp uniform convexity and smoothness inequalities for trace norms*, Invent. Math. **115** (1994), 463–482.

[Ba] V. Bangert, *Analytische Eigenschaften konvexer Funktionen auf Riemannschen Mannigfaltigkeiten* (German), J. Reine Angew. Math. **307/308** (1979), 309–324.

[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.

[BH] M. R. Bridson and A. Haefliger, Metric spaces of non-positive curvature, Springer-Verlag, Berlin, 1999.

[BBI] D. Burago, Yu. Burago and S. Ivanov, A course in metric geometry, American Mathematical Society, Providence, RI, 2001.

[Bu] H. Busemann, *Spaces with non-positive curvature*, Acta Math. **80** (1948), 259–310.

[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.

[Eg] D. Egloff, *Uniform Finsler Hadamard manifolds*, Ann. Inst. H. Poincaré Phys. Théor. **66** (1997), 323–357.

[GW1] R. E. Greene and H. Wu, *On the subharmonicity and plurisubharmonicity of geodesically convex functions*, Indiana Univ. Math. J. **22** (1972/73), 641–653.

[GW2] R. E. Greene and H. Wu, *Integrals of subharmonic functions on manifolds of nonnegative curvature*, Invent. Math. **27** (1974), 265–298.

[Jo] F. John, *Extremum problems with inequalities as subsidiary conditions*, Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

[LV1] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, to appear in Ann. of Math.

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

[Ly] A. Lytchak, *Differentiation in metric spaces*, St. Petersburg Math. J. **16** (2005), 1017–1041.

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

[Oh1] S. Ohta, *Regularity of harmonic functions in Cheeger-type Sobolev spaces*, Ann. Global Anal. Geom. **26** (2004), 397–410.

[Oh2] S. Ohta, *Convexities of metric spaces*, Geom. Dedicata **125** (2007), 225–250.

[Oh3] S. Ohta, *Markov type of Alexandrov spaces of nonnegative curvature*, to appear in Mathematika.

[Oh4] S. Ohta, *Gradient flows on Wasserstein spaces over compact Alexandrov spaces*, to appear in Amer. J. Math.

[Oh5] S. Ohta, *Extending Lipschitz and Hölder maps between metric spaces*, to appear in Positivity.

[Oh6] S. Ohta, *Finsler interpolation inequalities*, preprint (2008).

[RR] S. T. Rachev and L. Rüschendorf, Mass transportation problems. Vol. I, Springer-Verlag, New York, 1998.

[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, *Curvature, Distance and volume in Finsler geometry*, unpublished preprint.

[Sh3] Z. Shen, Lectures on Finsler geometry, World Scientific Publishing Co., Singapore, 2001.

[St1] K.-T. Sturm, *On the geometry of metric measure spaces*, Acta Math. **196** (2006), 65–131.

[St2] 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, 2008.

[WX] B. Y. Wu and Y. L. Xin, *Comparison theorems in Finsler geometry and their applications*, Math. Ann. **337** (2007), 177–196.