

On the curvature and heat flow on Hamiltonian systems

Shin-ichi OHTA*

Abstract

We develop the differential geometric and geometric analytic studies of Hamiltonian systems. Key ingredients are the curvature operator, the weighted Laplacian, and the associated Riccati equation. We prove appropriate generalizations of the Bochner–Weitzenböck formula and Laplacian comparison theorem, and study the heat flow.

MSC (2010): 53C21 (Primary); 58J35, 49Q20 (Secondary)

Keywords: Hamiltonian, curvature, comparison theorem, heat flow

Contents

1	Introduction	2
2	Preliminaries	3
2.1	Lagrangians	4
2.2	Hamiltonians	5
2.3	Finsler manifolds	6
3	Curvature for Hamiltonians	8
4	Laplacian	11
4.1	Gradient vectors, Laplacian and Hessian	11
4.2	Energy and harmonic functions	15
4.3	Riccati equation	16
4.4	Bochner–Weitzenböck formula	17
5	Examples	18
5.1	Finsler Hamiltonians	18
5.2	Natural mechanical Hamiltonians	19
5.3	Convex deformations of Finsler Hamiltonians	20
5.4	Homogeneous deformations of Finsler Hamiltonians	22

*Department of Mathematics, Kyoto University, Kyoto 606-8502, Japan (sohta@math.kyoto-u.ac.jp); Supported in part by the Grant-in-Aid for Young Scientists (B) 23740048.

6	Laplacian comparison theorem	22
6.1	Laplacian comparison	22
6.2	Measure contraction property	24
7	Optimal transports and functional inequalities	25
7.1	Optimal transport theory	25
7.2	Functional inequalities	27
8	Heat flow	30
8.1	Heat flow as gradient flow of energy	30
8.2	Gradient flow of the relative entropy	36
A	Appendix	39
A.1	How to construct canonical frames	39
A.2	Coordinate representation	41

1 Introduction

The aim of this article is to apply the recently developed technique in Finsler geometry to the study of Hamiltonian systems. A Finsler manifold carries a (Minkowski) norm on each tangent space. Although Finsler manifolds form a much wider class than Riemannian manifolds, the notion of curvature makes sense and we can consider various comparison theorems similarly to the Riemannian case (see, e.g., [Sh1]). Especially, the weighted Ricci curvature introduced by the author [Oh4] has fruitful applications including the curvature-dimension condition ([Oh4]), Laplacian comparison theorem for the natural nonlinear Laplacian ([OS1]), Bochner–Weitzenböck formula and gradient estimates ([OS3]), and generalizations of the Cheeger–Gromoll splitting theorem ([Oh6]). To be precise, the weighted Ricci curvature is defined for a pair consisting of a Finsler manifold and a measure on it, and our Laplacian depends on the choice of the measure (see Subsections 2.3, 4.1 for details).

Then, it is natural to expect that the theory of curvature can be applied beyond Finsler manifolds, and a class of manifolds M endowed with Lagrangians \mathbf{L} or, equivalently, Hamiltonians \mathbf{H} is a natural choice. In fact, on the one hand, we know that Agrachev and Gamkrelidze [AG] (see also [Agr2]) have developed the theory of curvature operator for Hamiltonian systems in connection with optimal control theory (see [Agr1] for a dynamical application). On the other hand, optimal transport theory (which is related to the curvature-dimension condition) for Lagrangian cost functions has been well investigated ([BB], [FF], [Vi]). Furthermore, Lee [Le2] recently showed a Riccati equation (see also [AL1], [AL2], [LLZ] for the sub-Riemannian case) as well as convexity estimates for entropy functionals along smooth optimal transports for general (time-dependent) Hamiltonian systems, by means of the curvature operator. His unified approach recovers both the curvature-dimension condition ($CD(K, \infty)$ and $CD(0, N)$ to be precise) for Riemannian or Finsler manifolds and various monotonicity formulas along flows in Riemannian metrics related to the Ricci flow.

Our Hamiltonian will always be time-independent and non-negative. Compared with the Finsler situation, the lack of the homogeneity causes many difficulties, for example, we need to take care of the difference between the Lagrangian and the Hamiltonian (they coincide as functions via the Legendre transform in the Finsler case). Nevertheless, by combining the Riccati equation and the technique in the Finsler case, we prove the Bochner–Weitzenböck formula (Theorem 4.4) and Laplacian comparison theorem (Theorem 6.2). We also obtain functional inequalities from the convexity of the relative entropy (Theorem 7.6). In these results, we use the weighted Ricci curvature derived from the curvature operator as well as the Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}}$ induced from the Hamiltonian \mathbf{H} and the reference measure \mathbf{m} on M . In general, this Laplacian is not only nonlinear ($\Delta_{\mathbf{m}}^{\mathbf{H}}(f + g) \neq \Delta_{\mathbf{m}}^{\mathbf{H}}f + \Delta_{\mathbf{m}}^{\mathbf{H}}g$) but also non-homogeneous ($\Delta_{\mathbf{m}}^{\mathbf{H}}(cf) \neq c\Delta_{\mathbf{m}}^{\mathbf{H}}f$ for $c \in \mathbb{R}$).

We also study the evolution equation $\partial_t u = \Delta_{\mathbf{m}}^{\mathbf{H}}u$ which can be thought of as the ‘heat equation’ in our context. In various settings including Finsler manifolds ([OS1]), the heat flow is regarded as gradient flow in the following two ways:

- (I) the gradient flow of the Dirichlet energy in the L^2 -space;
- (II) the gradient flow of the relative entropy in the L^2 -Wasserstein space.

The identification of these two strategies was recently established for general metric measure spaces satisfying the curvature-dimension condition ([AGS3]). Although a Hamiltonian does not induce a distance function, we verify the analogue of the former approach (I) in Theorem 8.4. Indeed, since our energy form $\mathcal{E}(u) = \int_M \mathbf{H}(du) d\mathbf{m}$ is a convex functional on the L^2 -space, the classical theory of Brézis et al applies. Interestingly, however, the latter strategy (II) leads to the different equation $\partial_t u = -\operatorname{div}_{\mathbf{m}}(u \nabla[-\log u])$ (Theorem 8.7). This shows that the identification of (I) and (II) is essentially due to the homogeneity of \mathbf{H} .

We explicitly calculate the curvature and Laplacian in the special class of Hamiltonians given by convex deformations of Finsler Hamiltonians (that is, $\mathbf{H} = h \circ F^*$, where h is a convex function and F^* is the dual of a Finsler structure $F : TM \rightarrow \mathbb{R}$; see Subsection 5.3). What is of particular interest is the p -homogeneous deformations ($h(t) = t^p/p$), which derives the Finsler analogue of p -Laplacians. However, we do not focus on this specific case because the aim of this work is to present the general framework. In the same spirit, we sometimes discuss under seemingly superfluous assumptions for the sake of technical simplicity.

Acknowledgments I am grateful to Professor Paul W. Y. Lee for his helpful comments on a preliminary version of the paper.

2 Preliminaries

Throughout the article, let M be a connected \mathcal{C}^∞ -manifold of dimension $n \geq 2$ without boundary. For a local coordinate system $(x^i)_{i=1}^n$ on an open set $U \subset M$, we will always consider the (fiber-wise linear) coordinates $(x^i; v^j)_{i,j=1}^n$ of the tangent bundle TU and

$(x^i; \alpha_j)_{i,j=1}^n$ of the co-tangent bundle T^*U given by

$$v = \sum_{j=1}^n v^j \partial_{x^j} = \sum_{j=1}^n v^j \frac{\partial}{\partial x^j}, \quad \alpha = \sum_{j=1}^n \alpha_j dx^j,$$

respectively. We will use the usual abbreviations such as

$$\mathbf{L}_{v^i} := \frac{\partial \mathbf{L}}{\partial v^i}, \quad \mathbf{H}_{x^i \alpha_j} := \frac{\partial^2 \mathbf{H}}{\partial \alpha_j \partial x^i}$$

for brevity, but only for the Lagrangian \mathbf{L} on TM and the Hamiltonian \mathbf{H} on T^*M .

2.1 Lagrangians

We consider only time-independent (autonomous) Lagrangians for simplicity. General time-dependent Lagrangians can be treated similarly to a great extent (see [AG], [Agr2], [Le2]).

Definition 2.1 (Lagrangians) A *Lagrangian* on M will mean a non-negative function $\mathbf{L} : TM \rightarrow [0, \infty)$ satisfying the following conditions:

- (1) $\mathbf{L} \in \mathcal{C}^1(TM) \cap \mathcal{C}^\infty(TM \setminus 0_{TM})$ and $\mathbf{L} \equiv 0$ on 0_{TM} , where $0_{TM} \subset TM$ denotes the zero section;
- (2) (*Super-linearity*) There are some complete Riemannian metric g of M and positive constant $C > 0$ such that $\mathbf{L}(v) \geq |v|_g - C$ for all $v \in TM$;
- (3) (*Strong convexity*) For any $x \in M$, \mathbf{L} has the positive-definite Hessian at every $v \in T_x M \setminus \{0\}$ (with respect to an arbitrary linear coordinate of $T_x M$).

Note that the strong convexity implies $\mathbf{L} > 0$ on $TM \setminus 0_{TM}$ and that the super-linearity (2) follows from (1) and (3) if M is compact. The reason why only the \mathcal{C}^1 -regularity is assumed on 0_{TM} is to include non-Riemannian Finsler metrics (see Remark 2.2(a)). Then the non-negativity $\mathbf{L} \geq 0$ and $\mathbf{L}|_{0_{TM}} \equiv 0$ are imposed to ensure that an action-minimizing curve $\eta : [0, T] \rightarrow M$ with $\dot{\eta}(0) \neq 0$ always enjoys $\dot{\eta} \neq 0$ on whole $[0, T]$ (due to the conservation of the Hamiltonian), thus $\dot{\eta}$ lives in $TM \setminus 0_{TM}$ where \mathbf{L} is \mathcal{C}^∞ . These additional conditions ($\mathbf{L} \geq 0$, $\mathbf{L}|_{0_{TM}} \equiv 0$) can be removed if $\mathbf{L} \in \mathcal{C}^2(TM)$ (as in [Le2]).

A \mathcal{C}^1 -curve $\eta : [0, T] \rightarrow M$ which is a critical point of the *action* $\int_0^T \mathbf{L}(\dot{\eta}) dt$ (among variations fixing the endpoints) is called an *action-minimizing curve* and satisfies the *Euler-Lagrange equation*

$$\mathbf{L}_{x^i}(\dot{\eta}) - \sum_{j=1}^n \{\mathbf{L}_{v^i x^j}(\dot{\eta}) \dot{\eta}^j + \mathbf{L}_{v^i v^j}(\dot{\eta}) \ddot{\eta}^j\} = 0 \quad \text{for all } i. \quad (2.1)$$

To be precise, η is either a constant curve or a \mathcal{C}^∞ -curve with $\dot{\eta} \neq 0$. Thus η is always \mathcal{C}^∞ and (2.1) makes sense (note that $\mathbf{L}_{x^i}(0) = 0$ since $\mathbf{L}|_{0_{TM}} \equiv 0$). Given any $v \in TM$, there is a unique action-minimizing curve $\eta_v : (-\varepsilon, \varepsilon) \rightarrow M$ with $\dot{\eta}_v(0) = v$ for sufficiently

small $\varepsilon > 0$. We say that (M, \mathbf{L}) is *forward complete* (resp. *complete*) if η_v is extended to the action-minimizing curve $\eta_v : [0, \infty) \rightarrow M$ (resp. $\eta_v : \mathbb{R} \rightarrow M$) for all $v \in TM$.

Let us summarize some fundamental remarks on the difference from the Riemannian or Finsler case, caused by the non-homogeneity of \mathbf{L} (see also Remark 2.3 below).

Remark 2.2 (a) If $\mathbf{L}(cv) = c^2\mathbf{L}(v)$ for all $v \in TM$ and $c > 0$, then $F(v) := \sqrt{2\mathbf{L}(v)}$ gives a Finsler metric (see Subsections 2.3, 5.1 below). Moreover, in this case, F comes from a Riemannian metric if and only if $F^2 \in \mathcal{C}^2(TM)$ (see [Sh1, Proposition 2.2]).

(b) An action-minimizing curve η does not necessarily have a constant speed (i.e., $\mathbf{L}(\dot{\eta})$ may not be constant). This is one of the reasons why the Hamiltonian (which is constant along the Hamiltonian flow) fits better to our consideration.

(c) The strong convexity does not imply the uniform (strict) convexity of \mathbf{L} even in a single tangent space T_xM . For instance, for $f(t) = |t|^p$ on \mathbb{R} with $p \in (1, \infty)$, we have $\lim_{t \rightarrow \infty} f''(t) = 0$ if $p < 2$ and $\lim_{t \downarrow 0} f''(t) = 0$ if $p > 2$. This is one of the major differences from the Finsler setting, in which the uniform convexity and smoothness are used in various analytic and geometric estimates (see [OS1, Sections 2, 3], [Oh3]).

2.2 Hamiltonians

Let \mathbf{L} be a Lagrangian as in Definition 2.1. The associated *Hamiltonian* is given by

$$\mathbf{H}(\alpha) := \sup_{v \in T_xM} \{\alpha(v) - \mathbf{L}(v)\}, \quad \alpha \in T_x^*M.$$

Choosing $v = 0$ ensures that $\mathbf{H} \geq 0$ as well as $\mathbf{H}|_{0_{T^*M}} \equiv 0$. Moreover, $\mathbf{H} > 0$ on $T^*M \setminus 0_{T^*M}$ and $\mathbf{H} \in \mathcal{C}^1(T^*M) \cap \mathcal{C}^\infty(T^*M \setminus 0_{T^*M})$. Given $\alpha \in T_x^*M$, by virtue of the strong convexity of \mathbf{L} , we can find a unique vector $v \in T_xM$ satisfying $\mathbf{H}(\alpha) = \alpha(v) - \mathbf{L}(v)$. Such a vector is denoted by $\boldsymbol{\tau}^*(\alpha)$ and called the *Legendre transform* of α , and the map $\boldsymbol{\tau}^* : T^*M \rightarrow TM$ is explicitly written as

$$\boldsymbol{\tau}^*(\alpha) = \sum_{i=1}^n \mathbf{H}_{\alpha_i}(\alpha) \partial_{x^i}. \quad (2.2)$$

The inverse map $\boldsymbol{\tau} : TM \rightarrow T^*M$ of $\boldsymbol{\tau}^*$ is similarly given by

$$\boldsymbol{\tau}(v) := \sum_{i=1}^n \mathbf{L}_{v^i}(v) dx^i.$$

Note that, for each $\alpha \neq 0$, $(\mathbf{H}_{\alpha_i \alpha_j}(\alpha))_{i,j=1}^n$ is the inverse matrix of $(\mathbf{L}_{v^i v^j}(\boldsymbol{\tau}^*(\alpha)))_{i,j=1}^n$ so that it is positive-definite.

Remark 2.3 (a) The transform $\boldsymbol{\tau}$ (or $\boldsymbol{\tau}^*$) is a linear operator if and only if \mathbf{L} comes from a Riemannian metric. Precisely, $\boldsymbol{\tau}(cv) = c\boldsymbol{\tau}(v)$ holds for all $c > 0$ and $v \in TM$ if and only if \mathbf{L} comes from a Finsler structure, and $\boldsymbol{\tau}(v+w) = \boldsymbol{\tau}(v) + \boldsymbol{\tau}(w)$ holds only in the Riemannian case. Moreover, in the Finsler case, $\boldsymbol{\tau}|_{T_xM}$ is differentiable at the origin if and only if $F|_{T_xM}$ comes from an inner product.

(b) In the Finsler setting, $\mathbf{H}(\boldsymbol{\tau}(v)) = \mathbf{L}(v)$ holds and the better estimate $\alpha(v) \leq 2\sqrt{\mathbf{L}(v)\mathbf{H}(\alpha)}$ is available.

The *Hamiltonian vector field* on T^*M is defined by

$$\vec{\mathbf{H}}(\alpha) := \sum_{i=1}^n \{\mathbf{H}_{\alpha_i}(\alpha) \partial_{x^i} - \mathbf{H}_{x^i}(\alpha) \partial_{\alpha_i}\} \in T_\alpha(T^*M).$$

Denote by $(\Phi_t(\alpha))_{t \in (-\varepsilon, \varepsilon)} \subset T^*M$ the corresponding *Hamiltonian flow*, i.e., $\Phi_0(\alpha) = \alpha$ and $\dot{\Phi}_t(\alpha) = \vec{\mathbf{H}}(\Phi_t(\alpha))$. The Hamiltonian is preserved along Φ_t , namely $\mathbf{H}(\Phi_t(\alpha)) = \mathbf{H}(\alpha)$ for all t .

A C^∞ -curve η solves the Euler–Lagrange equation (2.1) if and only if $\tau(\dot{\eta})$ verifies $\Phi_t(\tau(\dot{\eta}(0))) = \tau(\dot{\eta}(t))$. That is to say, the Hamiltonian flow coincides with the flow on TM generated from the Euler–Lagrange equation via the Legendre transform. Thus the forward completeness (resp. completeness) of \mathbf{L} is equivalent to that of Φ_t , namely the existence of $(\Phi_t)_{t \geq 0}$ (resp. $(\Phi_t)_{t \in \mathbb{R}}$) on whole T^*M .

Let us denote by $\omega = \sum_{i=1}^n d\alpha_i \wedge dx^i$ the canonical *symplectic form* on T^*M . Recall that

$$\omega(d\Phi_t(v), d\Phi_t(w)) = \omega(v, w) \quad \text{for all } v, w \in T_\alpha(T^*M), t \in (-\varepsilon, \varepsilon). \quad (2.3)$$

2.3 Finsler manifolds

This subsection is devoted to a concise review on the special class of Finsler manifolds, where we have a clearer understanding of curvature and heat flow. We refer to [BCS] and [Sh2] for basics of Finsler geometry.

Definition 2.4 (Finsler structures) We say that a function $F : TM \rightarrow [0, \infty)$ is a *C^∞ -Finsler structure* of M if it satisfies the following:

- (1) $F \in C^\infty(TM \setminus 0_{TM})$;
- (2) (*Positive 1-homogeneity*) $F(cv) = cF(v)$ for all $v \in TM$ and $c > 0$;
- (3) (*Strong convexity*) For any $v \in TM \setminus 0_{TM}$, the $(n \times n)$ -symmetric matrix

$$g_{ij}(v) := \frac{1}{2} \frac{\partial^2 (F^2)}{\partial v^i \partial v^j}(v) \quad (2.4)$$

is positive-definite.

We do not assume the *absolute homogeneity* $F(-v) = F(v)$ in general. Thanks to the positive homogeneity, the Lagrangian $\mathbf{L}^F := F^2/2$ and the corresponding Hamiltonian \mathbf{H}^F for instance satisfy

$$2\mathbf{L}^F(v) = \sum_{i,j=1}^n g_{ij}(v) v^i v^j, \quad \mathbf{H}^F(\alpha) = \frac{1}{2} F^*(\alpha)^2 = \mathbf{L}^F(\tau^*(\alpha))$$

for all $v \in T_x M$ and $\alpha \in T_x^* M$, where $F^*|_{T_x^* M}$ is the dual (Minkowski) norm of $F|_{T_x M}$ given by

$$F^*(\alpha) := \sup_{v \in T_x M, F(v) \leq 1} \alpha(v).$$

Euler's theorem on homogeneous functions is a key tool (see [BCS, Theorem 1.2.1]). In the context of general Lagrangians, the lack of this basic tool causes many differences from the Finsler setting.

Similarly to the Riemannian case, one can regard the Euler–Lagrange equation as the *geodesic equation* with respect to the distance structure induced from F . Precisely, the *distance function* $d_F : M \times M \rightarrow [0, \infty)$ associated with F is naturally introduced by

$$d_F(x, y) := \inf_{\eta} \int_0^1 F(\dot{\eta}) dt,$$

where the infimum is taken over all piece-wise \mathcal{C}^1 -curves $\eta : [0, 1] \rightarrow M$ with $\eta(0) = x$ and $\eta(1) = y$. Note that d_F can be nonsymmetric (i.e., $d_F(y, x) \neq d_F(x, y)$) since F is only positively homogeneous. Then a curve η solves the Euler–Lagrange equation if and only if it is a *geodesic* in the sense that it is locally d_F -minimizing and of constant speed (i.e., $F(\dot{\eta})$ is constant).

In the Finsler world, corresponding to the sectional curvature of Riemannian manifolds is the *flag curvature* $\mathbf{K}(v, w)$ for linearly independent vectors $v, w \in T_x M$. We remark that, different from the Riemannian case, $\mathbf{K}(v, w)$ depends not only on the plane $v \wedge w$ spanned by v and w (the *flag*), but also on the choice of v in it (the *flagpole*).

We know a useful interpretation of the flag curvature, due to Shen ([Sh2, §6.2]), as follows. Fix $v \in T_x M \setminus \{0\}$ and extend it to a \mathcal{C}^∞ -vector field V on a neighborhood U of x (i.e., $V(x) = v$) in such a way that all integral curves of V are geodesic (this is always possible, whereas the choice of V is not unique). By the strong convexity, V induces the Riemannian structure g_V on U via (2.4) as

$$g_V \left(\sum_{i=1}^n a_i \partial_{x^i}, \sum_{j=1}^n b_j \partial_{x^j} \right) := \sum_{i,j=1}^n a_i b_j g_{ij}(V). \quad (2.5)$$

Then, for $w \in T_x M$ which is not co-linear with v , the sectional curvature of $v \wedge w$ with respect to g_V coincides with $\mathbf{K}(v, w)$ (independent of the choice of V). This remarkable fact shows the usefulness of V and g_V as above in the study of Finsler manifolds from the Riemannian geometric viewpoint.

For a unit vector $v \in T_x M \cap F^{-1}(1)$, the *Ricci curvature* $\text{Ric}^F(v)$ is defined as the trace of $\mathbf{K}(v, \cdot)$ with respect to g_v (defined similarly to (2.5)). Thus $\text{Ric}^F(v)$ coincides with the Ricci curvature of v with respect to the Riemannian structure g_V as in the previous paragraph. We also set $\text{Ric}^F(cv) := c^2 \text{Ric}^F(v)$ for $c \geq 0$.

Now we fix a positive \mathcal{C}^∞ -measure \mathbf{m} on M and modify Ric^F into the *weighted Ricci curvature* Ric_N^F which was introduced in [Oh4] inspired by the theory of weighted Riemannian manifolds.

Definition 2.5 (Weighted Ricci curvature) Given $v \in T_x M \setminus \{0\}$, extend it to a \mathcal{C}^∞ -vector field V on a neighborhood U of x such that all integral curves of V are geodesic. Using the volume measure vol_V of g_V , we decompose \mathbf{m} as $\mathbf{m} = e^{-\psi} \text{vol}_V$ on U . Let η be the geodesic with $\dot{\eta}(0) = v$. Then we define, for $N \in (n, \infty)$,

$$\text{Ric}_N^F(v) := \text{Ric}^F(v) + (\psi \circ \eta)''(0) - \frac{(\psi \circ \eta)'(0)^2}{N - n}.$$

As the limits, define

$$\begin{aligned} \operatorname{Ric}_\infty^F(v) &:= \operatorname{Ric}^F(v) + (\psi \circ \eta)''(0), \\ \operatorname{Ric}_n^F(v) &:= \begin{cases} \operatorname{Ric}^F(v) + (\psi \circ \eta)''(0) & \text{if } (\psi \circ \eta)'(0) = 0, \\ -\infty & \text{if } (\psi \circ \eta)'(0) \neq 0. \end{cases} \end{aligned}$$

We also set $\operatorname{Ric}_N^F := 0$ on 0_{TM} for all $N \in [n, \infty]$.

It is easily seen that $\operatorname{Ric}_N^F(v)$ is well-defined (independent of the choice of V) and $\operatorname{Ric}_N^F(cv) = c^2 \operatorname{Ric}_N^F(v)$ for all $c > 0$. In the Riemannian case, $\operatorname{Ric}_\infty^g$ is the famous *Bakry–Émery tensor* and Ric_N^g with $N < \infty$ was introduced by Qian (see [BE], [Qi]).

Remark 2.6 On a Riemannian manifold (M, g) , if we choose the volume measure as the reference measure \mathbf{m} , then $\operatorname{Ric}_N^g = \operatorname{Ric}^g$ for all $N \in [n, \infty]$. On a general Finsler manifold, however, there does not necessarily exist a measure \mathbf{m} with $\operatorname{Ric}_n^F > -\infty$ (see [Oh5]). This means that there is no nice reference measure in general, so that it is natural to start from an arbitrary measure.

It was demonstrated in [Oh4] that (M, d_F, \mathbf{m}) satisfies Lott, Sturm and Villani’s *curvature-dimension condition* $\operatorname{CD}(K, N)$ if and only if $\operatorname{Ric}_N^F(v) \geq KF(v)^2$ for all $v \in TM$. Roughly speaking, in $\operatorname{CD}(K, N)$, K acts as the lower bound of the Ricci curvature and N is regarded as the upper bound of the dimension. This characterization extends the Riemannian one (by [CMS], [vRS], [St1], [St2], [LV1], [LV2]), and has many analytic and geometric applications via the general theory of curvature-dimension condition (developed in [St1], [St2], [LV1], [LV2], [Vi]). Moreover, the Laplacian and heat flow on (M, F, \mathbf{m}) (both are nonlinear except for the Riemannian case) were studied in [OS1], [OS2] and [OS3], where we have shown the Bochner–Weitzenböck formula and the Bakry–Émery and Li–Yau gradient estimates among others. See also [Oh6] for a further application to generalizations of the Cheeger–Gromoll splitting theorem.

3 Curvature for Hamiltonians

In [AG], Agrachev and Gamkrelidze introduced the curvature operator for Hamiltonian systems. We recall their definition along the line of [Le2] (see also [Le1]). Although our Lagrangian is assumed to be time-independent and non-negative, the original definition is concerned with general time-dependent Lagrangians.

Remark 3.1 According to [AD], the construction of curvature in [AG] is equivalent to those in the independent works [Gr] and [Fo] (see also Acknowledgment in the preprint version arXiv:1205.1442v6 of [Le2]). We refer to [AD] and the references therein for details and some other related works.

Fix $\alpha \in T_x^*M \setminus \{0\}$ and put $\alpha(t) := \Phi_t(\alpha)$ for $t \in (-\varepsilon, \varepsilon)$ throughout this section. For each $t \in (-\varepsilon, \varepsilon)$, let us consider

$$\begin{aligned} \mathcal{V}_{\alpha(t)} &:= \operatorname{span}\{\partial_{\alpha_i} \mid i = 1, \dots, n\} \subset T_{\alpha(t)}(T^*M), \\ \mathcal{J}_\alpha^t &:= (d\Phi_t)^{-1}(\mathcal{V}_{\alpha(t)}) \subset T_\alpha(T^*M). \end{aligned}$$

The curve $(\mathcal{J}_\alpha^t)_{t \in (-\varepsilon, \varepsilon)}$ in the n -dimensional Grassmannian manifold in $T(T^*M)$ is called a *Jacobi curve*, and represents how T^*M is distorted along the Hamiltonian flow Φ_t .

Fix $\xi = \sum_{i=1}^n \xi^i \partial_{\alpha_i} \in \mathcal{V}_\alpha$. Choose an arbitrary smooth curve $\boldsymbol{\xi}(t) = \sum_{i=1}^n \xi^i(t) \partial_{\alpha_i} \in \mathcal{V}_{\alpha(t)}$ with $\boldsymbol{\xi}(0) = \xi$, and put $e_t := (d\Phi_t)^{-1}(\boldsymbol{\xi}(t)) \in \mathcal{J}_\alpha^t$. Since e_t lives in the same linear space $T_\alpha(T^*M)$ for all t , we can define $\dot{e}_t \in T_\alpha(T^*M)$ by differentiating the coefficients of e_t (with respect to any linear coordinate of $T_\alpha(T^*M)$). Recall that ω denotes the canonical symplectic form on T^*M .

Lemma 3.2 *We have*

$$\omega(\dot{e}_0, e_0) = \sum_{i,j=1}^n \xi^i \xi^j \mathbf{H}_{\alpha_i \alpha_j}(\alpha).$$

In particular, $\omega(\dot{e}_0, e_0)$ is independent of the choice of the curve $\boldsymbol{\xi}$ and $\langle \xi, \xi \rangle_\alpha := \omega(\dot{e}_0, e_0)$ defines an inner product of \mathcal{V}_α .

Proof. Since $e_0 = \xi \in \mathcal{V}_\alpha$ is in the vertical part of $T_\alpha(T^*M)$, it is sufficient to calculate only the horizontal part of \dot{e}_0 . By differentiating $d\Phi_t(e_t) = \boldsymbol{\xi}(t)$ at $t = 0$ and noticing $\dot{\Phi}_t = \vec{\mathbf{H}} \circ \Phi_t$ and $e_0 = \xi$, we have

$$\dot{e}_0 + \sum_{i,j=1}^n \xi^j \{ \mathbf{H}_{\alpha_i \alpha_j}(\alpha) \partial_{x^i} - \mathbf{H}_{x^i \alpha_j}(\alpha) \partial_{\alpha_i} \} = \sum_{i=1}^n \dot{\xi}^i(0) \partial_{\alpha_i}.$$

Hence the horizontal part of \dot{e}_0 is $-\sum_{i,j=1}^n \xi^j \mathbf{H}_{\alpha_i \alpha_j}(\alpha) \partial_{x^i}$ and we obtain the claim. \square

Remark 3.3 The above inner product should be compared with (2.4) and (2.5) in the Finsler case. To be precise, $g_{\tau^*(\alpha)}$ of $T_x M$ coincides with $\langle \cdot, \cdot \rangle_\alpha$ of \mathcal{V}_α in Lemma 3.2 via the Legendre transform and the canonical identification between $T_x^* M$ and \mathcal{V}_α .

Definition 3.4 (Canonical frames) If a family of smooth curves $e_i^t \in \mathcal{J}_\alpha^t$ ($i = 1, \dots, n$) satisfies

$$\omega(\dot{e}_i^t, e_j^t) = \delta_{ij}, \quad \dot{e}_i^t \in \mathcal{J}_\alpha^t \quad \text{for all } t \in (-\varepsilon, \varepsilon), \quad (3.1)$$

then we call the family $(e_i^t; \dot{e}_j^t)_{i,j=1}^n$ (referred simply by $(e_i^t)_{i=1}^n$ henceforth) a *canonical frame* along α .

Indeed, the first condition in (3.1) ensures that $(e_i^t)_{i=1}^n$ spans \mathcal{J}_α^t and $(\dot{e}_i^t; \dot{e}_j^t)_{i,j=1}^n$ spans $T_\alpha(T^*M)$. We also remark that

$$\omega(\dot{e}_i^t, e_j^t) = \langle d\Phi_t(e_i^t), d\Phi_t(e_j^t) \rangle_{\alpha(t)}$$

by (2.3) and Lemma 3.2. In Appendix A.1, we will review how to construct a canonical frame from an orthonormal basis with respect to $\langle \cdot, \cdot \rangle_{\alpha(t)}$ along the recipe in [Le2].

Fix a canonical frame $E^t = (e_i^t)_{i=1}^n$ along α .

Lemma 3.5 *We have $\omega(e_i^t, e_j^t) = 0$ for all i, j and t . In particular, $\omega(e_i^t, \mathcal{J}_\alpha^t) = 0$ for all i and t .*

Proof. This follows from $\omega(e_i^0, e_j^0) = 0$ (since $e_i^0 \in \mathcal{V}_\alpha$) and

$$\partial_t[\omega(e_i^t, e_j^t)] = \omega(\dot{e}_i^t, e_j^t) + \omega(e_i^t, \dot{e}_j^t) = \delta_{ij} - \delta_{ji} = 0.$$

□

Hence \mathcal{J}_α^t is a Lagrangian subspace of $T_\alpha(T^*M)$ with respect to ω . The next lemma yields the uniqueness of a canonical frame up to an orthogonal transformation.

Lemma 3.6 (see [Le2, Proposition 3.1]) *Let $\bar{E}^t = (\bar{e}_i^t)_{i=1}^n$ be another canonical frame along α . Then there exists an orthogonal matrix $O = (O_{ij})_{i,j=1}^n$ such that $\bar{E}^t = OE^t$, i.e., $\bar{e}_i^t = \sum_{j=1}^n O_{ij}e_j^t$ for all i and t . In particular, $\bar{E}^t = E^t$ for all t if $\bar{E}^0 = E^0$.*

Proof. For each t , the matrix O^t given by $\bar{E}^t = O^t E^t$ is orthogonal since

$$I_n = \omega(\dot{\bar{E}}^t, \bar{E}^t) = O^t \omega(\dot{E}^t, E^t) (O^t)^\top = O^t (O^t)^\top,$$

where we used Lemma 3.5 and O^\top is the transpose of O . It also follows from Lemma 3.5 that

$$0 = \omega(\ddot{\bar{E}}^t, \bar{E}^t) = 2\dot{O}^t \omega(\dot{E}^t, E^t) (O^t)^{-1} = 2\dot{O}^t (O^t)^{-1}.$$

Therefore $\dot{O}^t \equiv 0$ and $O = O^0$ is our desired matrix. □

Definition 3.7 (Curvature operator) For each $t \in (-\varepsilon, \varepsilon)$, the *curvature operator* $\mathbf{R}_\alpha^t : \mathcal{J}_\alpha^t \rightarrow \mathcal{J}_\alpha^t$ is defined by

$$\mathbf{R}_\alpha^t(e_i^t) := -\ddot{e}_i^t.$$

Note that \mathbf{R}_α^t is a linear operator and is independent of the choice of the canonical frame $(e_i^t)_{i=1}^n$ thanks to Lemma 3.6. In Appendix A.2, we explicitly calculate the curvature operator in coordinates along [Le2]. The following property of \mathbf{R}_α^t shows that the definition of \mathbf{R}_α^t can be reduced to the case of $t = 0$.

Lemma 3.8 (see [Le2, (25)]) *For any t , we have $\mathbf{R}_{\alpha(t)}^0 = d\Phi_t \circ \mathbf{R}_\alpha^t \circ (d\Phi_t)^{-1}$.*

Proof. Fix t and observe that

$$\tilde{e}_i^s := d\Phi_t(e_i^{t+s}) \in d\Phi_t(\mathcal{J}_\alpha^{t+s}) = (d\Phi_s)^{-1}(\mathcal{V}_{\alpha(t+s)}) = \mathcal{J}_{\alpha(t)}^s$$

gives a canonical frame along $\alpha(t+s)$. Thus we obtain

$$\mathbf{R}_{\alpha(t)}^0(\tilde{e}_i^0) = -\ddot{\tilde{e}}_i^0 = -d\Phi_t(\ddot{e}_i^t) = d\Phi_t(\mathbf{R}_\alpha^t(e_i^t)) = d\Phi_t \circ \mathbf{R}_\alpha^t \circ (d\Phi_t)^{-1}(\tilde{e}_i^0).$$

□

Lemma 3.9 *The curvature operator \mathbf{R}_α^t is symmetric in the sense that*

$$\langle d\Phi_t(e_i^t), d\Phi_t(\mathbf{R}_\alpha^t(e_j^t)) \rangle_{\alpha(t)} = \langle d\Phi_t(\mathbf{R}_\alpha^t(e_i^t)), d\Phi_t(e_j^t) \rangle_{\alpha(t)}$$

for all i, j and t .

Proof. By Lemma 3.8, it suffices to see the claim at $t = 0$. We deduce from Lemma 3.5 that $\omega(e_i^t, \ddot{e}_j^t) \equiv 0$, and then $\omega(\dot{e}_i^t, \dot{e}_j^t) \equiv 0$ since $\omega(e_i^t, \dot{e}_j^t) \equiv -\delta_{ij}$. Thus we have

$$\langle e_i^0, \ddot{e}_j^0 \rangle_\alpha = \frac{1}{2} \{ \omega(\dot{e}_i^0, \ddot{e}_j^0) + \omega(\ddot{e}_j^0, \dot{e}_i^0) \} = \frac{1}{2} \{ \omega(\dot{e}_i^0, \ddot{e}_j^0) - \omega(\ddot{e}_j^0, \dot{e}_i^0) \} = \omega(\dot{e}_i^0, \ddot{e}_j^0),$$

and $\omega(\dot{e}_i^0, \ddot{e}_j^0) = -\omega(\ddot{e}_i^0, \dot{e}_j^0) = \omega(\dot{e}_j^0, \ddot{e}_i^0)$ shows the claim. \square

Definition 3.10 (Ricci curvature) Define the *Ricci curvature* $\text{Ric}^{\mathbf{H}}(\alpha) \in \mathbb{R}$ as the trace of $\mathbf{R}_\alpha^0 : \mathcal{V}_\alpha \rightarrow \mathcal{V}_\alpha$ with respect to the inner product $\langle \cdot, \cdot \rangle_\alpha$ given in Lemma 3.2. We also set $\text{Ric}^{\mathbf{H}} := 0$ on 0_{T^*M} .

We remark that setting $\text{Ric}^{\mathbf{H}}(0) = 0$ is reasonable since then α is constant. The weighted version can be introduced similarly to the Finsler case as follows (recall Definition 2.5).

Definition 3.11 (Weighted Ricci curvature) We fix a positive \mathcal{C}^∞ -measure \mathbf{m} on M . Along the action-minimizing curve $\eta(t) := \pi_M(\alpha(t)) \in M$ ($\pi_M : T^*M \rightarrow M$ is the canonical projection), decompose \mathbf{m} as $\mathbf{m} = e^{-\psi} \text{vol}_{\dot{\eta}}$, where $\text{vol}_{\dot{\eta}}$ is the volume form of $\mathbf{L}_{v^i v^j}(\dot{\eta})$ given by

$$\text{vol}_{\dot{\eta}} = \sqrt{\det[\mathbf{L}_{v^i v^j}(\dot{\eta})]} dx^1 \cdots dx^n$$

on η . Then the *weighted Ricci curvature* is defined by

$$\text{Ric}_N^{\mathbf{H}}(\alpha) := \text{Ric}^{\mathbf{H}}(\alpha) + (\psi \circ \eta)''(0) - \frac{(\psi \circ \eta)'(0)^2}{N - n}$$

for $N \in (n, \infty)$, and

$$\begin{aligned} \text{Ric}_\infty^{\mathbf{H}}(v) &:= \text{Ric}^{\mathbf{H}}(v) + (\psi \circ \eta)''(0), \\ \text{Ric}_n^{\mathbf{H}}(v) &:= \begin{cases} \text{Ric}^{\mathbf{H}}(v) + (\psi \circ \eta)''(0) & \text{if } (\psi \circ \eta)'(0) = 0, \\ -\infty & \text{if } (\psi \circ \eta)'(0) \neq 0. \end{cases} \end{aligned}$$

We also set $\text{Ric}_N^{\mathbf{H}} := 0$ on 0_{T^*M} for all $N \in [n, \infty]$.

4 Laplacian

In this section, we introduce the natural nonlinear Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}}$ associated with the Hamiltonian \mathbf{H} and the reference measure \mathbf{m} on M in a similar way to the Finsler case (see [OS1]). It will turn out that $\Delta_{\mathbf{m}}^{\mathbf{H}}$ coincides with the Laplacian $\Delta^{\mathbf{H}}$ studied in [Le2] (see also [AL2] for the sub-Riemannian case) up to a difference depending on \mathbf{m} .

4.1 Gradient vectors, Laplacian and Hessian

For a differentiable function u on M , we call

$$\nabla u(x) := \tau^*(du_x) = \sum_{i=1}^n \mathbf{H}_{\alpha_i}(du_x) \partial_{x^i} \in T_x M$$

the *gradient vector* of u at $x \in M$. For an open set $U \subset M$, we will use two kinds of *Sobolev spaces*:

$$\begin{aligned} H^1(U; \mathbf{L}) &:= \{u \in L^2(U; \mathfrak{m}) \mid \text{weakly differentiable, } \mathbf{L}(\nabla u) \in L^1(U; \mathfrak{m})\}, \\ H^1(U; \mathbf{H}) &:= \{u \in L^2(U; \mathfrak{m}) \mid \text{weakly differentiable, } \mathbf{H}(du) \in L^1(U; \mathfrak{m})\}. \end{aligned}$$

Clearly they coincide in the Riemannian or Finsler case. We also introduce $H_{\text{loc}}^1(U; \mathbf{L})$ and $H_{\text{loc}}^1(U; \mathbf{H})$ similarly. Note that $\mathcal{C}_c^\infty(U) \subset H^1(U; \mathbf{L}) \cap H^1(U; \mathbf{H})$ and $\mathcal{C}^\infty(U) \subset H_{\text{loc}}^1(U; \mathbf{L}) \cap H_{\text{loc}}^1(U; \mathbf{H})$. We remark that $H^1(U; \mathbf{L})$ is not necessarily a linear space because of the non-homogeneity of \mathbf{L} (same for $H^1(U; \mathbf{H})$, $H_{\text{loc}}^1(U; \mathbf{L})$ and $H_{\text{loc}}^1(U; \mathbf{H})$). We know only that $H^1(U; \mathbf{H})$ (resp. $H_{\text{loc}}^1(U; \mathbf{H})$) is a convex subset of $L^2(U; \mathfrak{m})$ (resp. $L_{\text{loc}}^2(U; \mathfrak{m})$) by the convexity of \mathbf{H} .

Associated with \mathfrak{m} , define the *divergence* of a (weakly) differentiable vector field V on M with $\mathbf{L}(V) \in L_{\text{loc}}^1(M; \mathfrak{m})$ in the weak form by

$$\int_M \phi \operatorname{div}_{\mathfrak{m}} V \, d\mathfrak{m} = - \int_M d\phi(V) \, d\mathfrak{m} \quad \text{for all } \phi \in \mathcal{C}_c^\infty(M).$$

Note that the right-hand side is well-defined since

$$|d\phi(V)| \leq \max\{\mathbf{H}(d\phi), \mathbf{H}(-d\phi)\} + \mathbf{L}(V).$$

If V is differentiable, then we can write down in coordinates as

$$\operatorname{div}_{\mathfrak{m}} V = \sum_{i=1}^n \left\{ \frac{\partial V^i}{\partial x^i} - V^i \frac{\partial \varsigma}{\partial x^i} \right\},$$

where $V = \sum_{i=1}^n V^i \partial_{x^i}$ and $\mathfrak{m} = e^{-\varsigma} dx^1 \cdots dx^n$.

Now we define the (distributional, weighted) *Laplacian* $\Delta_{\mathfrak{m}}^{\mathbf{H}}$ acting on functions $u \in H_{\text{loc}}^1(M; \mathbf{L})$ by $\Delta_{\mathfrak{m}}^{\mathbf{H}} u := \operatorname{div}_{\mathfrak{m}}(\nabla u)$, namely

$$\int_M \phi \Delta_{\mathfrak{m}}^{\mathbf{H}} u \, d\mathfrak{m} = - \int_M d\phi(\nabla u) \, d\mathfrak{m} \quad \text{for all } \phi \in \mathcal{C}_c^\infty(M).$$

Note that our Laplacian is a negative operator in the sense that

$$\int_M u \Delta_{\mathfrak{m}}^{\mathbf{H}} u \, d\mathfrak{m} = - \int_M du(\nabla u) \, d\mathfrak{m} = - \int_M \{\mathbf{H}(du) + \mathbf{L}(\nabla u)\} \, d\mathfrak{m} \leq 0$$

for all $u \in \mathcal{C}_c^\infty(M)$ and equality holds if and only if u is constant. We remark that, even if $u \in \mathcal{C}^\infty(M)$, ∇u may not be differentiable at points where ∇u vanishes (Remark 2.3(a)). On the set where $\nabla u \neq 0$, we can calculate

$$\begin{aligned} \Delta_{\mathfrak{m}}^{\mathbf{H}} u &= \operatorname{div}_{\mathfrak{m}} \left(\sum_{i=1}^n \mathbf{H}_{\alpha_i}(du) \partial_{x^i} \right) \\ &= \sum_{i=1}^n \left\{ \mathbf{H}_{\alpha_i x^i}(du) + \sum_{j=1}^n \mathbf{H}_{\alpha_i \alpha_j}(du) \frac{\partial^2 u}{\partial x^i \partial x^j} - \mathbf{H}_{\alpha_i}(du) \frac{\partial \varsigma}{\partial x^i} \right\}. \end{aligned} \quad (4.1)$$

The Laplacian studied in [Le2] can be regarded as an unweighted version of $\Delta_{\mathbf{m}}^{\mathbf{H}}$. Let us recall the definition in [Le2, §4]. Given $u \in \mathcal{C}^\infty(M)$, the image of the derivative $d(du)_x : T_x M \rightarrow T_{du_x}(T^*M)$ of $du : M \rightarrow T^*M$ is

$$P_x = \text{span} \left\{ \partial_{x^i} + \sum_{j=1}^n \frac{\partial^2 u}{\partial x^i \partial x^j}(x) \partial_{\alpha_j} \mid i = 1, \dots, n \right\}$$

(which is a Lagrangian subspace with respect to ω). We fix $x \in M$ with $du_x \neq 0$ and shall identify P_x with the graph of a linear map via a canonical frame $(e_i^t)_{i=1}^n$ along $\alpha(t) := \Phi_t(du_x)$. To be precise, we decompose as $T_{du_x}(T^*M) = \mathcal{V}_{du_x} \times \mathcal{H}_{du_x}$, where

$$\mathcal{H}_{du_x} := \text{span}\{\dot{e}_i^0 \mid i = 1, \dots, n\}$$

and is independent of the choice of a canonical frame by Lemma 3.6. Choose a coordinate around x such that $\mathbf{H}_{\alpha_i \alpha_j}(du_x) = \delta_{ij}$ for all i, j , and take the canonical frame $(e_i^t)_{i=1}^n$ with $e_i^0 = \partial_{\alpha_i}$. (We prefer this coordinate than a simpler one with $\mathbf{H}_{\alpha_i \alpha_j} \circ \alpha \equiv \delta_{ij}$ for the sake of visibility of the structure of the calculation.) Then we have, by Proposition A.2, (A.2) and (A.3),

$$\begin{aligned} \dot{e}_i^0 &= \sum_{j=1}^n (O_{ij}^t \bar{e}_j^t)'|_{t=0} = \frac{1}{2} \sum_{j=1}^n \Omega_{ij}^0 \partial_{\alpha_j} + \dot{\bar{e}}_i^0 \\ &= -\partial_{x^i} + \frac{1}{2} \sum_{j=1}^n \{(\mathbf{H}_{x^i \alpha_j} + \mathbf{H}_{x^j \alpha_i})(du_x) + \dot{\xi}_j^i(0) + \dot{\xi}_i^j(0)\} \partial_{\alpha_j}. \end{aligned}$$

Since $\sum_{k,l=1}^n \mathbf{H}_{\alpha_k \alpha_l}(\alpha(t)) \xi_i^k(t) \xi_j^l(t) = \delta_{ij}$ for all t and $\mathbf{H}_{\alpha_k \alpha_l}(du_x) = \delta_{kl}$, we have

$$\dot{\xi}_j^i(0) + \dot{\xi}_i^j(0) + [\mathbf{H}_{\alpha_i \alpha_j} \circ \alpha]'(0) = 0.$$

Hence

$$\dot{e}_i^0 = -\partial_{x^i} + \frac{1}{2} \sum_{j=1}^n \{(\mathbf{H}_{x^i \alpha_j} + \mathbf{H}_{x^j \alpha_i})(du_x) - [\mathbf{H}_{\alpha_i \alpha_j} \circ \alpha]'(0)\} \partial_{\alpha_j},$$

so that P_x is the graph of the linear map sending $-\dot{e}_i^0 \in \mathcal{H}_{du_x}$ to

$$\frac{1}{2} \sum_{j=1}^n \left\{ (\mathbf{H}_{x^i \alpha_j} + \mathbf{H}_{x^j \alpha_i})(du_x) - [\mathbf{H}_{\alpha_i \alpha_j} \circ \alpha]'(0) + 2 \frac{\partial^2 u}{\partial x^i \partial x^j}(x) \right\} \partial_{\alpha_j}. \quad (4.2)$$

The negative of this map is called the *Hessian* $\text{Hess}^{\mathbf{H}}u(x) : \mathcal{H}_{du_x} \rightarrow \mathcal{V}_{du_x}$ of u at x , and its trace

$$\Delta^{\mathbf{H}}u(x) := \sum_{i=1}^n \left\{ \mathbf{H}_{x^i \alpha_i}(du_x) - \frac{1}{2} [\mathbf{H}_{\alpha_i \alpha_i} \circ \alpha]'(0) + \frac{\partial^2 u}{(\partial x^i)^2}(x) \right\} \quad (4.3)$$

with respect to the canonical frame is called the Laplacian in [Le2].

To compare $\Delta_{\mathbf{m}}^{\mathbf{H}}u$ with $\Delta^{\mathbf{H}}u$, similarly to Definition 3.11, let us decompose \mathbf{m} along the action-minimizing curve $\eta(t) := \pi_M(\alpha(t))$ as

$$\mathbf{m} = e^{-\psi} \sqrt{\det[\mathbf{L}_{v^i v^j}(\dot{\eta})]} dx^1 \cdots dx^n.$$

Then we observe from (4.1) that, since $\sum_{i=1}^n \mathbf{H}_{\alpha_i}(du_x)\partial_{x^i} = \dot{\eta}(0)$,

$$\Delta_{\mathbf{m}}^{\mathbf{H}}u(x) = \sum_{i=1}^n \left\{ \mathbf{H}_{\alpha_i x^i}(du_x) + \frac{\partial^2 u}{(\partial x^i)^2}(x) \right\} - \left[\psi \circ \eta - \frac{1}{2} \log(\det[\mathbf{L}_{v^i v^j} \circ \dot{\eta}]) \right]'(0).$$

Note that

$$\begin{aligned} \left[\log(\det[\mathbf{L}_{v^i v^j} \circ \dot{\eta}]) \right]'(0) &= \left[\det[\mathbf{L}_{v^i v^j} \circ \dot{\eta}] \right]'(0) = - \left[\det[\mathbf{H}_{\alpha_i \alpha_j} \circ \boldsymbol{\alpha}] \right]'(0) \\ &= - \sum_{i=1}^n [\mathbf{H}_{\alpha_i \alpha_i} \circ \boldsymbol{\alpha}]'(0). \end{aligned}$$

Therefore we have

$$\Delta_{\mathbf{m}}^{\mathbf{H}}u(x) = \Delta^{\mathbf{H}}u(x) - (\psi \circ \eta)'(0), \quad (4.4)$$

so that $\Delta_{\mathbf{m}}^{\mathbf{H}}u$ and $\Delta^{\mathbf{H}}u$ coincide up to the term depending on the weight function ψ . In other words, $\Delta^{\mathbf{H}}$ can be interpreted as the weighted Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}}$ with respect to a measure \mathbf{m} such that ψ is constant along any action-minimizing curve η , whereas such a measure does not necessarily exist even on a Finsler manifold (Remark 2.6).

Remark 4.1 It can be checked by hand that the above Hessian coincides with $\nabla^2 u \in T^*M \otimes TM$ used in [OS3] to study the Bochner–Weitzenböck formula in the Finsler setting. To see this, recall from [OS3, Lemma 2.3] that

$$\nabla^2 u(\partial_{x^i}|_x) = \sum_{j=1}^n \left\{ \frac{\partial^2 u}{\partial x^i \partial x^j} - \sum_{k=1}^n \Gamma_{ij}^k(\nabla u) \frac{\partial u}{\partial x^k} \right\} (x) \partial_{x^j}|_x,$$

where $g_{ij}(\nabla u(x)) = \delta_{ij}$ is still assumed. We calculate by using the notations in [OS3] as, at x ,

$$\begin{aligned} &\sum_{k=1}^n \Gamma_{ij}^k(\nabla u) \frac{\partial u}{\partial x^k} \\ &= \frac{1}{2} \sum_{k=1}^n \left\{ \left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right) + \sum_{l=1}^n \frac{\partial g_{ij}}{\partial v^l} N_k^l \right\} (\nabla u) \frac{\partial u}{\partial x^k} \\ &= -\frac{1}{2} \sum_{k=1}^n \left(\frac{\partial g^{jk}}{\partial x^i} + \frac{\partial g^{ik}}{\partial x^j} - \frac{\partial g^{ij}}{\partial x^k} \right) (du) \frac{\partial u}{\partial x^k} + \frac{1}{2} \sum_{l=1}^n \frac{\partial g_{ij}}{\partial v^l} (\nabla u) G^l(\nabla u), \end{aligned}$$

where $(g^{ij}(du))$ is the inverse matrix of $(g_{ij}(\nabla u))$. It follows from the homogeneity that

$$\mathbf{H}_{x^i \alpha_j}(du_x) = \sum_{k=1}^n \mathbf{H}_{x^i \alpha_j \alpha_k}(du_x) \frac{\partial u}{\partial x^k}(x) = \sum_{k=1}^n \frac{\partial g^{jk}}{\partial x^i}(du_x) \frac{\partial u}{\partial x^k}(x).$$

Moreover, since $\dot{\eta}(0) = \nabla u(x)$, we have

$$[\mathbf{H}_{\alpha_i \alpha_j} \circ \boldsymbol{\alpha}]'(0) = \sum_{k=1}^n \left\{ \frac{\partial g^{ij}}{\partial x^k}(du_x) \frac{\partial u}{\partial x^k}(x) - \frac{\partial g_{ij}}{\partial v^k}(\dot{\eta}(0)) \dot{\eta}^k(0) \right\}.$$

Consequently, the geodesic equation $\ddot{\eta}^k + G^k(\dot{\eta}) \equiv 0$ shows that

$$\nabla^2 u(\partial_{x^i}|_x) = \frac{1}{2} \sum_{j=1}^n \left\{ 2 \frac{\partial^2 u}{\partial x^i \partial x^j}(x) + (\mathbf{H}_{x^i \alpha_j} + \mathbf{H}_{x^j \alpha_i})(du_x) - [\mathbf{H}_{\alpha_i \alpha_j} \circ \boldsymbol{\alpha}]'(0) \right\} \partial_{x^j}|_x.$$

Compare this with (4.2).

4.2 Energy and harmonic functions

Let $U \subset M$ be an open set. Define the *energy functional* $\mathcal{E}_U : H_{\text{loc}}^1(U; \mathbf{H}) \longrightarrow [0, \infty]$ by

$$\mathcal{E}_U(u) := \int_U \mathbf{H}(du) \, d\mathbf{m}.$$

Lemma 4.2 *The energy functional \mathcal{E}_U is lower semi-continuous on $L^2(U; \mathbf{m})$. Namely, for any sequence $\{u_i\}_{i \in \mathbb{N}} \subset L^2(U; \mathbf{m}) \cap H_{\text{loc}}^1(U; \mathbf{H})$ converging to $u \in L^2(U; \mathbf{m}) \cap H_{\text{loc}}^1(U; \mathbf{H})$ in $L^2(U; \mathbf{m})$, it holds*

$$\mathcal{E}_U(u) \leq \liminf_{i \rightarrow \infty} \mathcal{E}_U(u_i).$$

Proof. Given $\varepsilon > 0$, let $W \subset U$ be a compact set such that $\mathcal{E}_W(u) \geq \mathcal{E}_U(u) - \varepsilon$ (including the case where both energies are infinite). We cover W with finitely many, mutually disjoint open sets $\{U_k\}$ (up to an \mathbf{m} -negligible set) such that each U_k is diffeomorphic to the unit ball in \mathbb{R}^n . Then Serrin's classical theorem ([Se], see also [FGM]) is applicable on each U_k to obtain

$$\mathcal{E}_W(u) = \sum_k \mathcal{E}_{U_k}(u) \leq \sum_k \liminf_{i \rightarrow \infty} \mathcal{E}_{U_k}(u_i) \leq \liminf_{i \rightarrow \infty} \mathcal{E}_U(u_i).$$

We complete the proof by letting $\varepsilon \rightarrow 0$. □

We say that $u \in H_{\text{loc}}^1(U; \mathbf{L})$ is *weakly harmonic* on U if $\Delta_{\mathbf{m}}^{\mathbf{H}} u \equiv 0$ on U in the weak sense, that is,

$$\int_U d\phi(\nabla u) \, d\mathbf{m} = 0 \quad \text{for all } \phi \in \mathcal{C}_c^\infty(U).$$

Lemma 4.3 *Suppose that $u + \phi \in H^1(U'; \mathbf{H})$ holds for any open subset $U' \subset U$, $u \in H^1(U'; \mathbf{H})$ and $\phi \in \mathcal{C}_c^\infty(U')$. Then a function $u \in H^1(U; \mathbf{H}) \cap H_{\text{loc}}^1(U; \mathbf{L})$ is weakly harmonic on U if and only if*

$$\mathcal{E}_U(u) = \inf\{\mathcal{E}_U(u + \phi) \mid \phi \in \mathcal{C}_c^\infty(U)\}.$$

Similarly, $u \in H_{\text{loc}}^1(U; \mathbf{H}) \cap H_{\text{loc}}^1(U; \mathbf{L})$ is weakly harmonic on U if and only if, for any relatively compact open set $U' \subset U$,

$$\mathcal{E}_{U'}(u) = \inf\{\mathcal{E}_{U'}(u + \phi) \mid \phi \in \mathcal{C}_c^\infty(U')\}.$$

Proof. The convexity of \mathbf{H} yields that, for any $t \in (0, 1)$,

$$\begin{aligned}\mathbf{H}(du) - \mathbf{H}(d(u - \phi)) &\leq \frac{\mathbf{H}(d(u + t\phi)) - \mathbf{H}(du)}{t} \leq \mathbf{H}(d(u + \phi)) - \mathbf{H}(du), \\ \mathbf{H}(du) - \mathbf{H}(d(u - \phi)) &\leq \frac{\mathbf{H}(du) - \mathbf{H}(d(u - t\phi))}{t} \leq \mathbf{H}(d(u + \phi)) - \mathbf{H}(du).\end{aligned}$$

Hence the dominated convergence theorem implies that

$$\partial_t[\mathcal{E}_{U'}(u + t\phi)]|_{t=0} = \int_{U'} d\phi(\nabla u) \, dm,$$

which completes the proof of the both assertions. \square

Note that $u \in H^1(U; \mathbf{H})$ was necessary to ensure $\mathcal{E}_U(u) < \infty$, while $u \in H_{\text{loc}}^1(U; \mathbf{L})$ was used to make $\Delta_{\mathbf{m}}^{\mathbf{H}}u$ well-defined. The hypothesis of the lemma is fulfilled by, for example, the p -homogeneous deformations of Finsler (or Riemannian) Hamiltonians (see Subsection 5.4). In that case our Laplacian coincides with the (weighted) p -Laplacian.

4.3 Riccati equation

In [Le2, §4], Lee showed a Riccati type equation with respect to the Laplacian $\Delta^{\mathbf{H}}$ in (4.3) (see also [AL1], [AL2] and [LLZ] for the sub-Riemannian case). We repeat his argument for completeness, and derive a generalization of the Bochner–Weitzenböck formula along the line of [OS3] in the next subsection.

Let $u \in C^\infty(M)$ and fix $x \in M$ with $du_x \neq 0$. On a small neighborhood U of x on which $du \neq 0$, let us consider the solution $(u_t)_{t \in (-\varepsilon, \varepsilon)} \subset C^\infty(U)$ to the *Hamilton–Jacobi equation*

$$\partial_t u_t + \mathbf{H}(du_t) = 0, \quad u_0 = u|_U. \quad (4.5)$$

A geometric (or dynamical) meaning of (4.5) is that, for each $y \in U$,

$$\Phi_t(du_y) = (du_t)_{\mathcal{T}_t(y)} \quad (4.6)$$

holds (as far as $\mathcal{T}_t(y) \in U$), where $t \mapsto \mathcal{T}_t(y)$ is the action-minimizing curve with $\partial_t[\mathcal{T}_t(y)]|_{t=0} = \nabla u(y)$.

Put $\eta(t) := \mathcal{T}_t(x)$ and let $(e_i^t)_{i=1}^n$ be a canonical frame along $\alpha(t) := (du_t)_{\eta(t)}$. Then, for each fixed τ ,

$$\tilde{e}_i^{t,\tau} := d\Phi_\tau(e_i^t) \in d\Phi_\tau(\mathcal{J}_{du_x}^t) = \mathcal{J}_{\alpha(\tau)}^{t-\tau}$$

gives a canonical frame along $\alpha(\tau + t)$ (similarly to the proof of Lemma 3.8). We write down the map $d(du)_x : T_x M \rightarrow T_{du_x}(T^*M)$ as, for each t ,

$$d(du)_x \circ d\pi_M(\dot{e}_i^0) = \sum_{j=1}^n \{A_{ij}(t)\dot{e}_j^t + B_{ij}(t)\dot{e}_j^t\}, \quad i = 1, \dots, n, \quad (4.7)$$

where $A(t) = (A_{ij}(t))$ and $B(t) = (B_{ij}(t))$ are $(n \times n)$ -matrices. Applying $d\Phi_t$ yields

$$d\Phi_t \circ d(du)_x \circ d\pi_M(\dot{e}_i^0) = \sum_{j=1}^n \{A_{ij}(t)\dot{e}_j^{t,t} + B_{ij}(t)\dot{e}_j^{t,t}\}. \quad (4.8)$$

On the one hand, since $\tilde{e}_j^{t,t} \in \mathcal{V}_{\alpha(t)}$, (4.8) implies

$$\begin{aligned} d\mathcal{T}_t \circ d\pi_M(\dot{e}_i^0) &= d(\pi_M \circ \Phi_t \circ du)_x \circ d\pi_M(\dot{e}_i^0) = d\pi_M \left(\sum_{j=1}^n \{A_{ij}(t)\tilde{e}_j^{t,t} + B_{ij}(t)\dot{\tilde{e}}_j^{t,t}\} \right) \\ &= \sum_{j=1}^n B_{ij}(t)d\pi_M(\dot{\tilde{e}}_j^{t,t}). \end{aligned}$$

On the other hand, differentiating (4.6) at $y = x$ gives $d\Phi_t \circ d(du)_x = d(du_t)_{\eta(t)} \circ (d\mathcal{T}_t)_x$. By combining these with (4.8), we have

$$\sum_{j=1}^n B_{ij}(t)d(du_t)_{\eta(t)} \circ d\pi_M(\dot{\tilde{e}}_j^{t,t}) = \sum_{j=1}^n \{A_{ij}(t)\tilde{e}_j^{t,t} + B_{ij}(t)\dot{\tilde{e}}_j^{t,t}\}.$$

This shows that $\text{Hess}^{\mathbf{H}}u_t(\eta(t)) = -B(t)^{-1}A(t)$ in the frame $(\tilde{e}_i^{t,t}; \dot{\tilde{e}}_j^{t,t})$.

By differentiating (4.7) in t , we obtain

$$0 = \sum_{j=1}^n \left[\left\{ \dot{A}_{ij}(t) - \sum_{k=1}^n B_{ik}(t)\mathbf{R}_{kj}(t) \right\} e_j^t + \{A_{ij}(t) + \dot{B}_{ij}(t)\}\dot{e}_j^t \right],$$

where we set $\mathbf{R}_{\alpha}^t(e_j^t) = \sum_{k=1}^n \mathbf{R}_{jk}(t)e_k^t$. This implies

$$\dot{B}(t) = -A(t), \quad \dot{A}(t) = B(t)\mathbf{R}(t).$$

We consequently obtain the *matrix Riccati equation* ([Le2, (31)])

$$\partial_t [\text{Hess}^{\mathbf{H}}u_t(\eta(t))] + [\text{Hess}^{\mathbf{H}}u_t(\eta(t))]^2 + \mathbf{R}(t) = 0. \quad (4.9)$$

Taking the trace yields, by the symmetry of $\text{Hess}^{\mathbf{H}}u_t$ (see (4.2)) and Lemma 3.8,

$$\partial_t [\Delta^{\mathbf{H}}u_t(\eta(t))] + \|\text{Hess}^{\mathbf{H}}u_t(\eta(t))\|_{\text{HS}(du_t)}^2 + \text{Ric}^{\mathbf{H}}((du_t)_{\eta(t)}) = 0, \quad (4.10)$$

where $\|\cdot\|_{\text{HS}(du_t)}$ denotes the *Hilbert–Schmidt norm* with respect to a canonical frame along $\alpha(t) = (du_t)_{\eta(t)}$.

4.4 Bochner–Weitzenböck formula

Taking the reference measure \mathbf{m} into account, we readily observe from (4.10) and (4.4) that

$$\partial_t [\Delta_{\mathbf{m}}^{\mathbf{H}}u_t(\eta(t))] + \|\text{Hess}^{\mathbf{H}}u_t(\eta(t))\|_{\text{HS}(du_t)}^2 + \text{Ric}_{\infty}^{\mathbf{H}}((du_t)_{\eta(t)}) = 0. \quad (4.11)$$

From this, similarly to the Finsler case, one can derive the *Bochner–Weitzenböck formula*.

Theorem 4.4 (Bochner–Weitzenböck formula) *Let $u \in C^{\infty}(M)$. At any $x \in M$ with $du_x \neq 0$, we have*

$$\Delta_{\mathbf{m}}^{du}(\mathbf{H}(du)) - d(\Delta_{\mathbf{m}}^{\mathbf{H}}u)(\nabla u) = \text{Ric}_{\infty}^{\mathbf{H}}(du) + \|\text{Hess}^{\mathbf{H}}u\|_{\text{HS}(du)}^2, \quad (4.12)$$

$$\Delta_{\mathbf{m}}^{du}(\mathbf{H}(du)) - d(\Delta_{\mathbf{m}}^{\mathbf{H}}u)(\nabla u) \geq \text{Ric}_N^{\mathbf{H}}(du) + \frac{(\Delta_{\mathbf{m}}^{\mathbf{H}}u)^2}{N} \quad (4.13)$$

for $N \in [n, \infty)$, where

$$\Delta_{\mathfrak{m}}^{du} f := \operatorname{div}_{\mathfrak{m}} \left(\sum_{i,j=1}^n \mathbf{H}_{\alpha_i \alpha_j} (du) \frac{\partial f}{\partial x^j} \partial_{x^i} \right).$$

Proof. Calculate the first term in (4.11) at $t = 0$ as

$$\partial_t [\Delta_{\mathfrak{m}}^{\mathbf{H}} u_t(\eta(t))] |_{t=0} = d(\Delta_{\mathfrak{m}}^{\mathbf{H}} u)_x(\dot{\eta}(0)) + \operatorname{div}_{\mathfrak{m}} \left(\sum_{i=1}^n \partial_t [\mathbf{H}_{\alpha_i}((du_t)_x)] |_{t=0} \partial_{x^i} \right).$$

Then (4.12) follows from $\dot{\eta}(0) = \nabla u(x)$ and (4.5) since

$$\partial_t [\mathbf{H}_{\alpha_i}((du_t)_x)] |_{t=0} = - \sum_{i,j=1}^n \mathbf{H}_{\alpha_i \alpha_j}((du)_x) \frac{\partial [\mathbf{H}(du)]}{\partial x^j}(x).$$

One can derive (4.13) from (4.12) in a standard way with the help of (4.4) (see the proof of [OS3, Theorem 3.3] for instance). \square

It is easily seen that $\Delta_{\mathfrak{m}}^{du} f$ is well-defined on $\{x \in M \mid du_x \neq 0\}$.

Remark 4.5 (a) The formulas in Theorem 4.4 coincide with those in [OS3, Theorem 3.3] on Finsler manifolds. We remark that, different from the Finsler setting (see [OS1, Lemma 2.4]), $\Delta_{\mathfrak{m}}^{du} u = \Delta_{\mathfrak{m}}^{\mathbf{H}} u$ does not necessarily hold due to the non-homogeneity of \mathbf{H} .

(b) In the Finsler case, one can treat functions with lower regularity by passing to the weak (integrated) formulations of (4.12) and (4.13) ([OS3, Theorem 3.6]). It is a more delicate issue for general Hamiltonians. Furthermore, we could obtain Bakry–Émery and Li–Yau gradient estimates in [OS3, §4] as applications of the (weak form of) Bochner–Weitzenböck formulas (see also [Oh6], [WX], [Xi] for further applications). In these proofs, however, we were essentially indebted to the homogeneity of the Finsler metric (Euler’s theorem [OS3, Theorem 2.2] to be precise) and it is unclear whether these gradient estimates can be generalized to general Hamiltonians.

5 Examples

This section is devoted to discussing several examples of Hamiltonians and calculating the curvature and Laplacian of them.

5.1 Finsler Hamiltonians

Let (M, F) be a Finsler manifold as in Subsection 2.3.

Example 5.1 Let $\mathbf{L}^F(v) := F(v)^2/2$ be the canonical Lagrangian induced from F , and \mathbf{H}^F be the associated Hamiltonian. Then the Hamiltonian curvature operator $\mathbf{R}_{\alpha}^0 : \mathcal{V}_{\alpha} \rightarrow \mathcal{V}_{\alpha}$ for $\alpha \in T_x^*M \setminus \{0\}$ coincides with the Finsler curvature operator $\mathbf{R}^F(\cdot, \boldsymbol{\tau}^*(\alpha))\boldsymbol{\tau}^*(\alpha) : T_x M \rightarrow T_x M$ (with the reference vector $\boldsymbol{\tau}^*(\alpha)$) by identifying \mathcal{V}_{α} and $T_x M$ (as in Remark 3.3). In particular, we have $\operatorname{Ric}^{\mathbf{H}}(\alpha) = \operatorname{Ric}^F(\boldsymbol{\tau}^*(\alpha))$.

One can verify this coincidence by the Riccati equation (4.9) (whose Finsler version can be found in [OS3, Lemmas 3.1, 3.2]) for instance (see also [Le2, §11]). Recall that there is a useful interpretation of the Finsler curvature by using Riemannian structures induced from vector fields satisfying a certain condition (see Subsection 2.3). For general Lagrangians, however, it seems difficult to obtain an analogous interpretation because even $2\mathbf{L}(v) = \sum_{i,j=1}^n \mathbf{L}_{v^i v^j}(v) v^i v^j$ fails.

5.2 Natural mechanical Hamiltonians

Let (M, g) be a Riemannian manifold and $Z \in \mathcal{C}^\infty(M)$.

Example 5.2 (Agrachev) The *natural mechanical Hamiltonian*

$$\mathbf{H}(\alpha) := \mathbf{H}^g(\alpha) + Z(x), \quad \alpha \in T_x^*M,$$

yields

$$\mathbf{R}_\alpha^0 = \mathbf{R}^g(\cdot, \boldsymbol{\tau}^*(\alpha)) \boldsymbol{\tau}^*(\alpha) + \text{Hess}^g Z(x) \quad (5.1)$$

by identifying \mathcal{V}_α on the LHS and $T_x M$ on the RHS as in Example 5.1, where $\boldsymbol{\tau}^*$ is the Legendre transform common to \mathbf{H}^g and \mathbf{H} . In particular, we have $\text{Ric}^{\mathbf{H}}(\alpha) = \text{Ric}^g(\boldsymbol{\tau}^*(\alpha)) + \Delta^g Z(x)$.

We remark that \mathbf{H} is allowed to be negative since it is \mathcal{C}^∞ on whole T^*M . Since Z is constant on each $T_x M$, we immediately find that $\mathbf{L}(v) = \mathbf{L}^g(v) - Z(x)$ for $v \in T_x M$, and that the Legendre transform of \mathbf{H} coincides with that of \mathbf{H}^g . One can also see that the Hessian and Laplacian are common to \mathbf{H}^g and \mathbf{H} . Then, in the Riccati equation (4.9), the term $\text{Hess}^g Z$ in (5.1) comes from the difference between the Hamilton–Jacobi equations (4.5) of \mathbf{H}^g and \mathbf{H} . To see this, given $u \in \mathcal{C}^\infty(M)$ and $x \in M$ with $du_x \neq 0$, let (u_t) and (\tilde{u}_t) be the solutions to (4.5) around x , with $u_0 = \tilde{u}_0 = u$, with respect to \mathbf{H}^g and \mathbf{H} , respectively. By comparing the matrix representations of $\text{Hess}^{\mathbf{H}} \tilde{u}_t(x)$ and $\text{Hess}^g u_t(x)$ as in (4.2), we have

$$\begin{aligned} & \partial_t[\text{Hess}^{\mathbf{H}} \tilde{u}_t(x)]|_{t=0} - \partial_t[\text{Hess}^g u_t(x)]|_{t=0} \\ &= -\frac{1}{2} \sum_{k=1}^n \{ \mathbf{H}_{x^i \alpha_j \alpha_k}^g + \mathbf{H}_{x^j \alpha_i \alpha_k}^g - \mathbf{H}_{\alpha_i \alpha_j x^k}^g \}(x) \frac{\partial Z}{\partial x^k}(x) - \frac{\partial^2 Z}{\partial x^i \partial x^j}(x) \\ &= -\frac{\partial^2 Z}{\partial x^i \partial x^j}(x) + \sum_{k=1}^n \Gamma_{ij}^k(x) \frac{\partial Z}{\partial x^k}(x) \\ &= -\text{Hess}^g Z(x). \end{aligned}$$

We used $\mathbf{H}_{\alpha_i \alpha_j \alpha_k} \equiv 0$ in the first equality. This yields (5.1) (with $\alpha = du_x$). One can alternatively show (5.1) by the direct computation in (A.4).

The weighted curvature can be treated similarly as follows. Let \mathbf{m} be a positive \mathcal{C}^∞ -measure on M . In the current situation, we can write down as $\mathbf{m} = e^{-\psi} \text{vol}_g$ globally. Since the Legendre transform is common to \mathbf{H}^g and \mathbf{H} , the weighted Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}}$ coincides

with Δ_m^g . Thus we find from (4.1) that, for (u_t) and (\tilde{u}_t) as above,

$$\begin{aligned} & \partial_t[\Delta_m^{\mathbf{H}}\tilde{u}_t(x)]|_{t=0} - \partial_t[\Delta_m^g u_t(x)]|_{t=0} \\ &= - \sum_{i,j=1}^n \left\{ \mathbf{H}_{\alpha_i x^i \alpha_j}^g \frac{\partial Z}{\partial x^j} + \mathbf{H}_{\alpha_i \alpha_j}^g \frac{\partial^2 Z}{\partial x^i \partial x^j} - \mathbf{H}_{\alpha_i \alpha_j}^g \frac{\partial \varsigma}{\partial x^i} \frac{\partial Z}{\partial x^j} \right\} (x) \\ &= -\Delta_m^g Z(x). \end{aligned}$$

We used Euler's theorem in the last line. Hence we have, by comparing (4.11) for \mathbf{H}^g and \mathbf{H} ,

$$\text{Ric}_{\infty}^{\mathbf{H}}(\alpha) = \text{Ric}_{\infty}^g(\tau^*(\alpha)) + \Delta_m^g Z(x).$$

Then, since $\mathbf{L}_{v^i v^j} = \mathbf{L}_{v^i v^j}^g$ and the difference $\text{Ric}_N^{\mathbf{H}}(\alpha) - \text{Ric}_{\infty}^{\mathbf{H}}(\alpha)$ involves only the first order derivatives of the weight function ψ , we immediately obtain

$$\text{Ric}_N^{\mathbf{H}}(\alpha) = \text{Ric}_N^g(\tau^*(\alpha)) + \Delta_m^g Z(x)$$

for $N \in [n, \infty)$ as well.

5.3 Convex deformations of Finsler Hamiltonians

Let (M, F) be a Finsler manifold and $h : \mathbb{R} \rightarrow [0, \infty)$ be a \mathcal{C}^∞ -function with $h(0) = h'(0) = 0$, $h'' > 0$ on $(0, \infty)$, as well as $\lim_{t \rightarrow \infty} h'(t) = \infty$. In particular, $h > 0$ on $(0, \infty)$ and $h' : [0, \infty) \rightarrow [0, \infty)$ is bijective. Now, let us consider the Hamiltonian $\mathbf{H}(\alpha) := h(F^*(\alpha))$. Since

$$\tau^*(\alpha) = \sum_{i=1}^n h'(F^*(\alpha)) \frac{\partial F^*}{\partial \alpha_i}(\alpha) \partial_{x^i} = \frac{h'(F^*(\alpha))}{F^*(\alpha)} \tau_F^*(\alpha)$$

on $T^*M \setminus 0_{T^*M}$, where τ_F^* is the Legendre transform with respect to F , we have

$$\Delta_m^{\mathbf{H}} u = \text{div}_m \left(\frac{h'(F^*(du))}{F^*(du)} \nabla^F u \right) = \frac{h'(F^*(du))}{F^*(du)} \Delta_m^F u + d \left[\frac{h'(F^*(du))}{F^*(du)} \right] (\nabla^F u) \quad (5.2)$$

on $\{x \in M \mid du_x \neq 0\}$. One can see this also by calculating (4.1). The calculation of the curvature is a little more involved, for that we go back to the definition of the curvature operator. The following proposition shows that the ratio of $\text{Ric}^{\mathbf{H}}(\alpha)$ to $\text{Ric}^F(\tau_F^*(\alpha))$ depends only on $F^*(\alpha)$ and h .

Proposition 5.3 *Let $\mathbf{H}(\alpha) = h(F^*(\alpha))$ be as above. Then we have*

$$\mathbf{R}_\alpha^0 = c(\alpha)^2 \tilde{\mathbf{R}}_\alpha^0, \quad \text{Ric}^{\mathbf{H}}(\alpha) = c(\alpha)^2 \text{Ric}^F(\tau_F^*(\alpha))$$

for any $\alpha \in T^*M \setminus 0_{T^*M}$, where $\tilde{\mathbf{R}}_\alpha^0$ denotes the curvature operator (in the sense of Definition 3.7) with respect to \mathbf{H}^F and we set $c(\alpha) := h'(F^*(\alpha))/F^*(\alpha)$. It also holds that $\text{Ric}_N^{\mathbf{H}}(\alpha) = c(\alpha)^2 \text{Ric}_N^F(\tau_F^*(\alpha))$ for any $N \in [n, \infty]$.

Proof. Put $c := c(\alpha)$ and $\tilde{\alpha}(t) := \Phi_t^F(\alpha)$ for $t \in (-\varepsilon, \varepsilon)$, and take a coordinate $(x^i)_{i=1}^n$ such that $dx^1 = F^*(\tilde{\alpha}(t))^{-1}\tilde{\alpha}(t)$ and $(\partial_{x^i})_{i=1}^n$ is orthonormal with respect to $g_{\tau_F^*(\alpha(t))}$ for all t . Then we can take a canonical frame $(\tilde{e}_i^t)_{i=1}^n$ along $\tilde{\alpha}(t)$ with respect to \mathbf{H}^F such that $\tilde{e}_1^t = (d\Phi_t^F)^{-1}(\partial_{\alpha_1}|_{\tilde{\alpha}(t)})$. (Start from $\tilde{e}_i^t := (d\Phi_t^F)^{-1}(\partial_{\alpha_i}|_{\tilde{\alpha}(t)})$ in Appendix A.1 and observe that $\Omega_{i1}^t \equiv 0$ for all i from $\tilde{e}_1^t \equiv 0$ and Lemma A.1, then we have $O_{11}^t \equiv 1$ and $O_{1i}^t = O_{i1}^t \equiv 0$ for $i \geq 2$ by decomposing the equation in Proposition A.2.)

Switching to \mathbf{H} , we first observe from

$$\vec{\mathbf{H}}(\tilde{\alpha}(t)) = \frac{h'(F^*(\tilde{\alpha}(t)))}{F^*(\tilde{\alpha}(t))} \vec{\mathbf{H}}^F(\tilde{\alpha}(t)) = c \vec{\mathbf{H}}^F(\tilde{\alpha}(t))$$

that $\Phi_t(\alpha) = \Phi_{ct}^F(\alpha)$. Put $\alpha(t) := \Phi_t(\alpha) = \tilde{\alpha}(ct)$ and note that, by $\mathbf{H}_{\alpha_i} = [h'(F^*)/F^*]\mathbf{H}_{\alpha_i}^F$, $\mathbf{H}_{\alpha_i\alpha_j}^F(\alpha(t)) = \delta_{ij}$, the choice of the coordinate and Euler's theorem,

$$\begin{aligned} \mathbf{H}_{\alpha_i\alpha_j}(\alpha(t)) &= c\delta_{ij} \quad \text{for } i \geq 2 \text{ or } j \geq 2, \\ \mathbf{H}_{\alpha_1\alpha_1}(\alpha(t)) &= c + \left[\frac{h''(F^*)F^* - h'(F^*)}{(F^*)^2} \frac{\partial F^*}{\partial \alpha_1} \mathbf{H}_{\alpha_1}^F \right] (\alpha(t)) \\ &= c + \left[\frac{h''(F^*)F^* - h'(F^*)}{(F^*)^2} \frac{F^*}{F^*} \frac{2\mathbf{H}^F}{F^*} \right] (\alpha(t)) \\ &= h''(F^*(\alpha)). \end{aligned}$$

Hence

$$e_1^t := \frac{1}{\sqrt{h''(F^*(\alpha))}} \tilde{e}_1^{ct}, \quad e_i^t := \frac{1}{\sqrt{c}} \tilde{e}_i^{ct} \quad \text{for } i \geq 2$$

give a canonical frame along $\alpha(t)$ with respect to \mathbf{H} . Now we have

$$\mathbf{R}_\alpha^t(e_1^t) = -\dot{e}_1^t = \frac{c^2}{\sqrt{h''(F^*(\alpha))}} \tilde{\mathbf{R}}_\alpha^{ct}(\tilde{e}_1^{ct}) = c^2 \tilde{\mathbf{R}}_\alpha^{ct}(e_1^t),$$

and similarly $\mathbf{R}_\alpha^t(e_i^t) = c^2 \tilde{\mathbf{R}}_\alpha^{ct}(e_i^t)$ for $i \geq 2$. These show $\mathbf{R}_\alpha^0 = c^2 \tilde{\mathbf{R}}_\alpha^0$ as well as $\text{Ric}^{\mathbf{H}}(\alpha) = c^2 \text{Ric}^F(\tau_F^*(\alpha))$.

The weighted version immediately follows from the above calculations. To be precise, denoting by $\tilde{\psi}$ and ψ the weight functions as in Definition 3.11 with respect to \mathbf{H}^F and \mathbf{H} , we find $\psi \circ \pi_M \circ \alpha(t) = \tilde{\psi} \circ \pi_M \circ \tilde{\alpha}(ct) + C$ with some constant C (depending on $F^*(\alpha)$ and h). This yields $\text{Ric}_N^{\mathbf{H}}(\alpha) = c^2 \text{Ric}_N^F(\tau_F^*(\alpha))$. \square

For instance, if $h(t) = (at)^2/2$ with some $a > 0$, then we have

$$\text{Ric}^{\mathbf{H}}(\alpha) = a^4 \text{Ric}^F(\tau_F^*(\alpha)) = a^2 \text{Ric}^F(\tau_F^*(\alpha)) \cdot \frac{\mathbf{H}(\alpha)}{\mathbf{H}^F(\alpha)}.$$

When we deform the Lagrangian as $\mathbf{L}(v) := h(F(v))$ instead of the Hamiltonian, we observe from

$$\tau(v) = \frac{h'(F(v))}{F(v)} \tau_F(v), \quad h(F(v)) + \mathbf{H}(\tau(v)) = [\tau(v)](v) = h'(F(v))F(v)$$

that $F^*(\tau(v)) = h'(F(v))$ and

$$\mathbf{H}(\alpha) = F^*(\alpha) \cdot (h')^{-1}(F^*(\alpha)) - h \circ (h')^{-1}(F^*(\alpha)).$$

Since the function $\bar{h}(t) := t(h')^{-1}(t) - h \circ (h')^{-1}(t)$ is convex (note that $\bar{h}' = (h')^{-1}$) and satisfies $\bar{h}(0) = \bar{h}'(0) = 0$, we can apply Proposition 5.3 and obtain

$$\mathrm{Ric}_N^{\mathbf{H}}(\alpha) = \left(\frac{(h')^{-1}(F^*(\alpha))}{F^*(\alpha)} \right)^2 \mathrm{Ric}_N^F(\tau_F^*(\alpha)).$$

5.4 Homogeneous deformations of Finsler Hamiltonians

Let (M, F) be a Finsler manifold again. One of the most important examples of convex deformations of F^* is the *p-homogeneous deformation*:

$$\mathbf{H}_p(\alpha) := \frac{F^*(\alpha)^p}{p}, \quad 1 < p < \infty.$$

The corresponding Lagrangian is $\mathbf{L}_q(v) := F(v)^q/q$, where $q = p/(p-1)$ is the dual exponent of p . By (5.2), \mathbf{H}_p leads to the (Finsler analogue of) famous *p-Laplacian*

$$\Delta_{\mathbf{m}}^{\mathbf{H}_p} u = \mathrm{div}_{\mathbf{m}}(F^*(du)^{p-2} \cdot \nabla^F u).$$

One can perform better analysis for $\Delta_{\mathbf{m}}^{\mathbf{H}_p}$ thanks to the homogeneity, although we do not pursue that direction in this paper (except for Example 8.8, we refer to recent works [Ke1], [Ke2] instead). Note also that $\mathrm{Ric}_N^{\mathbf{H}_p}(\alpha) = F^*(\alpha)^{2(p-2)} \mathrm{Ric}_N^F(\tau_F^*(\alpha))$ by Proposition 5.3. This suggests that the behavior of $\Delta_{\mathbf{m}}^{\mathbf{H}_p} u$ depends on how large $F^*(du)$ is.

6 Laplacian comparison theorem

We return to a general Hamiltonian \mathbf{H} on M and show a generalization of the Laplacian comparison theorem associated with the weighted Ricci curvature $\mathrm{Ric}_N^{\mathbf{H}}$ (suggested in [Le2, Remark 4.4], see also [AL2] for a related work in the sub-Riemannian setting). We also refer to [OS1, Theorem 5.2] for the Finsler case and to [Gi] for the more general case of metric measure spaces enjoying the curvature-dimension condition. Our Laplacian comparison theorem would be compared with the Bonnet–Myers type theorem of Agrachev and Gamkrelidze ([AG, Theorem in §4], [Agr2, Theorem 2.1]). Throughout the section, we assume the forward completeness of the Hamiltonian flow Φ_t .

6.1 Laplacian comparison

Fix $z \in M$ and $c > 0$. We introduce the (dual of) *exponential map of scale c*, $\exp_z^c : T_z^*M \setminus \{0\} \rightarrow M$, by $\exp_z^c(\alpha) := \eta(a^{-1})$ where $a > 0$ is chosen as $\mathbf{H}(a\alpha) = c$ and η is the action-minimizing curve with $\dot{\eta}(0) = \tau^*(a\alpha)$. In other words, $\exp_z^c(\alpha) = \pi_M(\Phi_{a^{-1}}(a\alpha))$. Set also $\exp_z^c(0) := z$.

Remark 6.1 In the Riemannian and Finsler cases, the map \exp_z^c coincides with the ordinary exponential map \exp_z for all $c > 0$ (via the Legendre transform as usual). For the p -homogeneous deformation \mathbf{H}_p of a Finsler metric as in Subsection 5.4, it holds that

$$\exp_z^{c'}(\alpha) = \exp_z^c \left(\left(\frac{c'}{c} \right)^{(p-2)/p} \alpha \right) \quad \text{for } c, c' > 0, \alpha \in T_z^*M$$

since $\mathbf{H}_p(a\alpha) = c'$ implies

$$\mathbf{H}_p \left(\left(\frac{c}{c'} \right)^{(p-1)/p} a \cdot \left(\frac{c'}{c} \right)^{(p-2)/p} \alpha \right) = \mathbf{H}_p \left(\left(\frac{c}{c'} \right)^{1/p} a\alpha \right) = \frac{c}{c'} \mathbf{H}_p(a\alpha) = c$$

and

$$\boldsymbol{\tau}^* \left(\left(\frac{c}{c'} \right)^{1/p} a\alpha \right) = \left(\frac{c}{c'} \right)^{(p-1)/p} \boldsymbol{\tau}^*(a\alpha).$$

Put

$$D_z^c := \{t\alpha \in T_z^*M \mid \mathbf{H}(\alpha) = c, 0 < t < T_\alpha\}, \quad \mathcal{D}_z^c := \exp_z^c(D_z^c),$$

where $T_\alpha \in (0, \infty]$ is the supremum of $T > 0$ for which $d\exp_z^c$ is invertible on $\{t\alpha\}_{t \in (0, T)}$ and, for all $t \in (0, T)$, there is no $\alpha' \in T_z^*M \setminus \{\alpha\}$ satisfying $\mathbf{H}(\alpha') = c$ as well as $\exp_z^c(t\alpha') = \exp_z^c(t\alpha)$. Then we define the one-form $\boldsymbol{\alpha}$ on \mathcal{D}_z^c by $\boldsymbol{\alpha}(\exp_z^c(t\alpha)) = \boldsymbol{\tau}(\dot{\eta}(t)) = \Phi_t(\alpha)$ for $t\alpha \in D_z^c$, where $\mathbf{H}(\alpha) = c$ and η is the action-minimizing curve with $\dot{\eta}(0) = \boldsymbol{\tau}^*(\alpha)$. Integrating along $\eta(t) = \exp_z^c(t\alpha)$ gives the function $u_z^c \in \mathcal{C}^\infty(\mathcal{D}_z^c)$ satisfying $\lim_{x \rightarrow z} u_z^c(x) = 0$ and $du_z^c = \boldsymbol{\alpha}$. Since $(u_z^c \circ \eta)'(t) = \boldsymbol{\alpha}(\dot{\eta}(t)) = c + \mathbf{L}(\dot{\eta}(t))$, we have

$$u_z^c(\eta(t)) = ct + \int_0^t \mathbf{L}(\dot{\eta}) ds.$$

Theorem 6.2 (Laplacian comparison) *Assume that $\text{Ric}_N^{\mathbf{H}}(\alpha) \geq K$ holds for some $K \in \mathbb{R}$, $N \in [n, \infty)$, and all $\alpha \in T^*M$ with $\mathbf{H}(\alpha) = c$. Then we have, for any $z \in M$, $\alpha \in T_z^*M$ with $\mathbf{H}(\alpha) = c$, and the action-minimizing curve $\eta : [0, T_\alpha] \rightarrow \overline{\mathcal{D}_z^c}$ with $\dot{\eta}(0) = \boldsymbol{\tau}^*(\alpha)$,*

$$\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t)) \leq \sqrt{NK} \cot \left(\sqrt{\frac{K}{N}} t \right) \quad \text{for all } t \in (0, T_\alpha) \quad (6.1)$$

if $K > 0$. It similarly holds that $\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t)) \leq \sqrt{-NK} \coth(\sqrt{-K/N} t)$ if $K < 0$, and $\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t)) \leq N/t$ if $K = 0$.

Proof. It follows from (4.11) with $u_t = u_z^c - ct$ and Theorem 4.4 that

$$\partial_t [\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t))] \leq -\text{Ric}_N^{\mathbf{H}}((du_z^c)_{\eta(t)}) - \frac{[\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t))]^2}{N}$$

for $t \in (0, T_\alpha)$. We have $\text{Ric}_N^{\mathbf{H}}((du_z^c)_{\eta(t)}) = \text{Ric}_N^{\mathbf{H}}(\boldsymbol{\alpha}(\eta(t))) \geq K$ by assumption, therefore

$$\partial_t [\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t))] \leq -K - \frac{[\Delta_{\mathfrak{m}}^{\mathbf{H}} u_z^c(\eta(t))]^2}{N}$$

for all $t \in (0, T_\alpha)$.

Put $u(t) := \Delta_{\mathbf{m}}^{\mathbf{H}} u_z^c(\eta(t))$ for brevity and compare it with $\tilde{u}(t) := N \mathbf{s}'_{K,N}(t) / \mathbf{s}_{K,N}(t)$, where

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

Observe that $\tilde{u}' = -K - \tilde{u}^2/N$ and $\lim_{t \downarrow 0} \mathbf{s}_{K,N}(t)^2 \tilde{u}(t) = 0$. Then the desired inequality $u \leq \tilde{u}$ follows from $\lim_{t \downarrow 0} \mathbf{s}_{K,N}(t)^2 u(t) = 0$ and the calculation (see [WW, (3.10)])

$$\begin{aligned} [\mathbf{s}_{K,N}^2(u - \tilde{u})]' &= 2\mathbf{s}_{K,N} \mathbf{s}'_{K,N}(u - \tilde{u}) + \mathbf{s}_{K,N}^2(u - \tilde{u})' \leq 2 \frac{\mathbf{s}_{K,N}^2 \tilde{u}}{N} (u - \tilde{u}) - \mathbf{s}_{K,N}^2 \frac{u^2 - \tilde{u}^2}{N} \\ &= -\frac{\mathbf{s}_{K,N}^2}{N} (u - \tilde{u})^2 \leq 0. \end{aligned}$$

Finally, $\lim_{t \downarrow 0} \mathbf{s}_{K,N}(t)^2 u(t) = 0$ can be seen as follows. Choose a local coordinate $(x^i)_{i=1}^n$ of a small neighborhood U of z such that every action-minimizing curve σ with $\sigma(0) = z$ and $\mathbf{H}(\boldsymbol{\tau}(\dot{\sigma})) \equiv c$ is represented as $\sigma(t) = (tb^i)_{i=1}^n$ with $\sum_{i=1}^n (b^i)^2 = 1$. Then it holds that $\nabla u_z^c(x) = r^{-1} \sum_{i=1}^n x^i \partial_{x^i}$ on $U \setminus \{z\}$, where $r = \sqrt{\sum_{i=1}^n (x^i)^2}$. Hence we find $\lim_{t \downarrow 0} t^2 u(t) = 0$ by

$$r^2 \sum_{i=1}^n \partial_{x^i} [r^{-1} x^i] = nr - \sum_{i=1}^n r^{-1} (x^i)^2 = (n-1)r \rightarrow 0$$

as $r \downarrow 0$. □

Remark 6.3 In the Riemannian and Finsler cases, $u_z^c / \sqrt{2c}$ coincides with the distance function d_z from z , and (6.1) is improved to

$$\Delta_{\mathbf{m}} d_z \leq \sqrt{(N-1) \frac{K}{2c}} \cot \left(\sqrt{\frac{1}{N-1} \frac{K}{2c}} d_z \right) \quad (6.2)$$

by decomposing $\Delta_{\mathbf{m}} d_z$ into the radial and spherical directions from z . It is unclear if a similar improvement can be done for general Hamiltonians (see also the Bonnet–Myers type theorem of Agrachev and Gamkrelidze). The point is that $u_z^c(\eta(t))$ is not proportional to t since $(u_z^c \circ \eta)'(t) = c + \mathbf{L}(\dot{\eta}(t))$ and $\mathbf{L}(\dot{\eta})$ is not necessarily constant. We also remark that even the stronger inequality (6.2) does not imply the curvature bound $\text{Ric}_N^{\mathbf{H}} \geq K$ (see [St2, Remark 5.6] for a simple example, and [Ju] for a related work on the gap between the measure contraction property and the curvature-dimension condition in Heisenberg groups).

6.2 Measure contraction property

A geometric counterpart to the Laplacian comparison is the *measure contraction property* (see [Oh1], [St2, §5] for the precise definition on metric measure spaces, and [AL1], [LLZ] for related works on sub-Riemannian manifolds). To state it, we fix $c > 0$, $z \in M$,

$\alpha \in T_z^*M$ and $\eta : [0, T_\alpha] \rightarrow M$ as in Theorem 6.2, and take a local coordinate $(x^i)_{i=1}^n$ of a neighborhood U of $\eta(T)$ with $T \in (0, T_\alpha)$ such that $\nabla u_z^c \equiv \partial_{x^1}$. We also introduce $\varsigma : U \rightarrow \mathbb{R}$ by $\mathbf{m} = e^\varsigma dx^1 \cdots dx^n$. Then the measure contraction property we consider is that the ratio

$$\frac{e^{\varsigma(\eta(t))}}{\mathbf{s}_{K,N}(t)^N}$$

is non-increasing in t . This is clearly equivalent to

$$(\varsigma \circ \eta)'(t) \leq N \frac{\mathbf{s}'_{K,N}(t)}{\mathbf{s}_{K,N}(t)}.$$

Since $\nabla u_z^c \equiv \partial_{x^1}$ and $\dot{\eta}(t) = \nabla u_z^c(\eta(t))$, we observe

$$\operatorname{div}_{\mathbf{m}}(\nabla u_z^c)(\eta(t)) = \frac{\partial \varsigma}{\partial x^1}(\eta(t)) = (\varsigma \circ \eta)'(t).$$

Therefore the above measure contraction property is equivalent to the Laplacian comparison (6.1):

$$\Delta_{\mathbf{m}}^{\mathbf{H}} u_z^c(\eta(t)) \leq N \frac{\mathbf{s}'_{K,N}(t)}{\mathbf{s}_{K,N}(t)}.$$

Remark 6.4 In [Le2, Theorems 2.2, 2.10], Lee showed convexity estimates of the relative and Rényi entropies along smooth optimal transports. These can be interpreted as generalizations of the curvature-dimension condition (recall the last paragraph in Section 2), and has applications to various monotonicity formulas along flows in Riemannian metrics related to the Ricci flow ([Le2, §2]). Precisely, this approach recovers the curvature-dimension conditions $\operatorname{CD}(K, \infty)$ and $\operatorname{CD}(0, N)$ for Riemannian or Finsler manifolds. The condition $\operatorname{CD}(K, N)$ with $K \neq 0$ and $N < \infty$ is more delicate and seems difficult to treat for general Hamiltonians for the same reasoning as Remark 6.3.

7 Optimal transports and functional inequalities

In this section, we briefly recall some properties of optimal transports measured by Lagrangian cost functions. Then, assuming that the relative entropy is convex along all optimal transports, we obtain functional inequalities along the lines of [OV], [LV2]. See the comprehensive reference [Vi] for optimal transport theory.

Let M be compact throughout the section for simplicity, and \mathbf{L} be our Lagrangian. For later use, we fix an auxiliary Riemannian metric g of M .

7.1 Optimal transport theory

For $x, y \in M$ and $T > 0$, we consider the *Lagrangian cost function* given by

$$c_T^{\mathbf{L}}(x, y) := \inf_{\eta} \int_0^T \mathbf{L}(\dot{\eta}) dt,$$

where the infimum is taken over all \mathcal{C}^1 -curves $\eta : [0, T] \rightarrow M$ with $\eta(0) = x$ and $\eta(T) = y$. Note that $c_T^{\mathbf{L}g}(x, y) = d_g(x, y)^2/(2T)$, where d_g is the distance function with respect to g .

Denote by $\mathcal{P}(M)$ the set of all Borel probability measures on M . Given $\mu, \nu \in \mathcal{P}(M)$, $\pi \in \mathcal{P}(M \times M)$ is called a *coupling* of μ and ν if its projections to the first and second arguments coincide with μ and ν , i.e., $\pi[A \times M] = \mu[A]$ and $\pi[M \times A] = \nu[A]$ for any Borel set $A \subset M$. Denote by $\Pi(\mu, \nu)$ the set of all couplings of μ and ν , which is nonempty since the product measure $\mu \times \nu$ is clearly a coupling of μ and ν . Then the *transport cost* from μ to ν (measured by \mathbf{L}) is defined as

$$C_T^{\mathbf{L}}(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{M \times M} c_T^{\mathbf{L}}(x, y) \pi(dx dy), \quad (7.1)$$

which is finite since M is compact. A coupling $\pi \in \Pi(\mu, \nu)$ achieving the infimum in (7.1) is called an *optimal coupling* of μ and ν . Under mild assumptions, there exists a unique optimal coupling whose support is drawn as the graph of some map $\mathcal{T} : M \rightarrow M$ (an *optimal transport* from μ to ν). This fundamental fact was first shown for the canonical Lagrangian on Euclidean spaces by Brenier [Bre], and extended to compact Riemannian manifolds by McCann [Mc], and then to general Lagrangians on noncompact spaces by Bernard and Buffoni [BB], Fathi and Figalli [FF] and Villani [Vi, Chapter 10].

Theorem 7.1 (see [Vi, Theorem 10.28]) *Take $\mu, \nu \in \mathcal{P}(M)$ such that $\mu \ll \mathfrak{m}$. Then there exists a semi-convex function φ (with respect to g) such that the map*

$$\mathcal{T}_t(x) := \pi_M(\Phi_t(d\varphi_x))$$

gives a unique optimal coupling $\pi = (\text{id}_M \times \mathcal{T}_T)_\# \mu$ of μ and ν , i.e., $(\mathcal{T}_T)_\# \mu = \nu$ and

$$C_T^{\mathbf{L}}(\mu, \nu) = \int_M c_T^{\mathbf{L}}(x, \mathcal{T}_T(x)) \mu(dx).$$

Denoted by $f_\# \mu$ was the push-forward measure of μ by a map f . Note that [Vi, Theorem 10.28] is certainly available by [Vi, Proposition 10.15, Theorem 10.26]. Observe also that, by [Vi, Proposition 10.15(ii), (10.20)],

$$d\varphi_x = -d[c_T^{\mathbf{L}}(\cdot, \mathcal{T}_T(x))]_x = \boldsymbol{\tau}(\dot{\eta}(0))$$

for μ -almost every x , where $\eta : [0, T] \rightarrow M$ is the unique action-minimizing curve from x to $\mathcal{T}_T(x)$. When we put $\mu_t := (\mathcal{T}_t)_\# \mu$, it clearly holds that $C_T^{\mathbf{L}}(\mu, \nu) = C_t^{\mathbf{L}}(\mu, \mu_t) + C_{T-t}^{\mathbf{L}}(\mu_t, \nu)$ for all $t \in (0, T)$. We have $\mu_t \ll \mathfrak{m}$ for every $t \in (0, T)$ (see [Vi, Theorem 8.7]). The semi-convexity of φ implies that φ is twice differentiable μ -almost everywhere, so that \mathcal{T}_t is differentiable μ -almost everywhere.

Remark 7.2 Lagrangians are assumed to be \mathcal{C}^2 on whole TM at some places in [Vi]. However, it causes no problem because our Lagrangian fails to be \mathcal{C}^2 only on the zero section 0_{TM} which is isolated in the Euler–Lagrange flow (recall the paragraph following Definition 2.1).

To state the *change of variables* (or the *Jacobian equation*) for the map \mathcal{T}_t , we define the *Jacobian determinant* of \mathcal{T}_t at $x \in M$ with respect to \mathbf{m} by

$$\mathbf{D}_{\mathbf{m}}[\mathcal{T}_t](x) := \lim_{r \downarrow 0} \frac{\mathbf{m}[\mathcal{T}_t(B^g(x, r))]}{\mathbf{m}[B^g(x, r)]}$$

if the limit exists, where $B^g(x, r)$ is the open ball of center x and radius r with respect to the Riemannian distance induced from our auxiliary Riemannian metric g . Note that, at points where \mathcal{T}_t is differentiable, $\mathbf{D}_{\mathbf{m}}[\mathcal{T}_t]$ does not depend on the choice of g .

Theorem 7.3 (Change of variables) *Let μ, ν and \mathcal{T}_t be as in Theorem 7.1, and put $\mu_t := (\mathcal{T}_t)_\# \mu = \rho_t \mathbf{m}$. Then we have, for any $t \in (0, T)$ and nonnegative measurable function $f : [0, \infty) \rightarrow [0, \infty)$ with $f(0) = 0$,*

$$\int_M f(\rho_t) d\mathbf{m} = \int_M f\left(\frac{\rho_0}{\mathbf{D}_{\mathbf{m}}[\mathcal{T}_t]}\right) \mathbf{D}_{\mathbf{m}}[\mathcal{T}_t] d\mathbf{m}. \quad (7.2)$$

If in addition $\nu \ll \mathbf{m}$, then (7.2) holds also at $t = T$.

Proof. This is a simple application of [Vi, Theorem 11.3]. Decompose \mathbf{m} as $\mathbf{m} = e^{-\psi} \text{vol}_g$. Applying [Vi, Theorem 11.3] to $F(x, s) = e^{-\psi(x)} f(se^{\psi(x)})$ yields

$$\int_M e^{-\psi} f(\rho_t) d\text{vol}_g = \int_M e^{-\psi(\mathcal{T}_t)} f\left(\frac{\rho_0 e^{-\psi}}{\mathbf{D}_{\mathbf{m}}^g[\mathcal{T}_t]} e^{\psi(\mathcal{T}_t)}\right) \mathbf{D}_{\mathbf{m}}^g[\mathcal{T}_t] d\text{vol}_g,$$

where $\mathbf{D}_{\mathbf{m}}^g[\mathcal{T}_t]$ denotes the Jacobian determinant with respect to vol_g . Together with $\mathbf{D}_{\mathbf{m}}[\mathcal{T}_t] = e^{\psi - \psi(\mathcal{T}_t)} \mathbf{D}_{\mathbf{m}}^g[\mathcal{T}_t]$, we have

$$\int_M f(\rho_t) d\mathbf{m} = \int_M f\left(\frac{\rho_0}{\mathbf{D}_{\mathbf{m}}[\mathcal{T}_t]}\right) \mathbf{D}_{\mathbf{m}}[\mathcal{T}_t] d\mathbf{m}.$$

(Precisely, one needs to separate the discussion into $M_\varphi := \{x \in M \mid d\varphi_x \neq 0\}$ and $M \setminus M_\varphi$. On $M \setminus M_\varphi$, \mathcal{T}_t is the identity map and $\mathbf{D}_{\mathbf{m}}[\mathcal{T}_t] = 1$ for all t , μ -almost everywhere.) \square

7.2 Functional inequalities

For $\mu \in \mathcal{P}(M)$, the *relative entropy* with respect to \mathbf{m} is defined by

$$\text{Ent}_{\mathbf{m}}(\mu) := \int_M \rho \log \rho d\mathbf{m}$$

if $\mu = \rho \mathbf{m} \ll \mathbf{m}$, and $\text{Ent}_{\mathbf{m}}(\mu) := \infty$ otherwise. Note that $\text{Ent}_{\mathbf{m}}(\mu) \geq -\log(\mathbf{m}[M])$ by Jensen's inequality.

Lee in [Le2, Theorem 2.2] showed that, along an optimal transport $(\mu_t)_{t \in [0, T]}$ as in Theorem 7.1 such that $\varphi \in \mathcal{C}^\infty(M)$, it holds that

$$\partial_t^2[\text{Ent}_{\mathbf{m}}(\mu_t)] \geq \int_M \text{Ric}_\infty^{\mathbf{H}}(\Phi_t(d\varphi_x)) \mu_0(dx) \quad \text{for all } t.$$

Therefore, if $\text{Ric}_\infty^{\mathbf{H}}(\alpha) \geq K\mathbf{H}(\alpha)$ for some $K \in \mathbb{R}$ and all $\alpha \in T^*M$, then we have

$$\text{Ent}_m(\mu_{(1-t)a+bt}) \leq (1-t)\text{Ent}_m(\mu_a) + t\text{Ent}_m(\mu_b) - \frac{K}{2}(1-t)t(b-a)^2 \int_M \mathbf{H}(d\varphi_x) \mu_0(dx) \quad (7.3)$$

for all $0 \leq a < b \leq T$ and $t \in [0, 1]$. We set

$$C_T^{\mathbf{H}}(\mu_0, \mu_1) := T \int_M \mathbf{H}(d\varphi_x) \mu_0(dx) = \int_0^T \int_M \mathbf{H}(\Phi_t(d\varphi_x)) \mu_0(dx) dt$$

for later use. Note that $C_T^{\mathbf{H}} = C_T^{\mathbf{L}}$ in the Riemannian and Finsler cases.

We say that Ent_m is K -convex for $K \in \mathbb{R}$ and $T > 0$ if (7.3) holds for all (not necessarily smooth) optimal transports $(\mu_t)_{t \in [0, T]}$ as in Theorem 7.1.

Example 7.4 (a) For Finsler manifolds, thanks to the homogeneity, the K -convexity for some $T > 0$ implies the K -convexity for all $T > 0$. In this case, the $2K$ -convexity of Ent_m with respect to $\mathbf{H}^F(\alpha) = F^*(\alpha)^2/2$ is the very definition of the curvature-dimension condition $\text{CD}(K, \infty)$ and is equivalent to $\text{Ric}_\infty^F(v) \geq KF(v)^2$ (recall Subsection 2.3).

(b) Assume that a Finsler manifold (M, F, μ) satisfies $\text{Ric}_\infty^F(v) \geq KF(v)^2$. Then, for the p -homogeneous deformation $\mathbf{H}_p(\alpha) = F^*(\alpha)^p/p$, we have

$$\text{Ric}_\infty^{\mathbf{H}_p}(\alpha) = F^*(\alpha)^{2(p-2)} \text{Ric}_\infty^F(\tau_F^*(\alpha)) \geq KF^*(\alpha)^{2(p-2)+2} = pK\mathbf{H}_p(\alpha)F^*(\alpha)^{p-2}.$$

To estimate $F^*(\alpha)$ for $K \neq 0$, observe that $TF^*(d\varphi) \leq \text{diam}_F M$ along all optimal transports $(\mu_t)_{t \in [0, T]}$, where $\text{diam}_F M := \sup_{x, y \in M} d_F(x, y)$ is the diameter of (M, F) . Therefore Ent_m is $(pK(\text{diam}_F M/T)^{p-2})$ -convex with respect to \mathbf{H}_p for $T > 0$ when (1) $K = 0$ and $p \in (1, \infty)$; or (2) $K > 0$ and $p \in (1, 2)$; or (3) $K < 0$ and $p \in (2, \infty)$.

The following proposition will be useful.

Proposition 7.5 (Directional derivatives of Ent_m) *Suppose that Ent_m is K -convex for some $K \in \mathbb{R}$ and $T > 0$, and take $\mu = \rho\mathbf{m} \in \mathcal{P}(M)$ such that $\rho \in H^1(M; \mathbf{H})$. Then we have, for any optimal transport $\mu_t = (\mathcal{I}_t)_\# \mu$, $t \in [0, T]$, as in Theorem 7.1,*

$$\liminf_{t \downarrow 0} \frac{\text{Ent}_m(\mu_t) - \text{Ent}_m(\mu)}{t} \geq \int_M \frac{d\rho}{\rho} (\nabla \varphi) d\mu. \quad (7.4)$$

Proof. It follows from Theorem 7.3 with $f(t) = t \log t$ that

$$\begin{aligned} \text{Ent}_m(\mu_t) &= \int_M \rho_t \log \rho_t d\mathbf{m} = \int_M \frac{\rho}{\mathbf{D}_m[\mathcal{I}_t]} \log \left(\frac{\rho}{\mathbf{D}_m[\mathcal{I}_t]} \right) \mathbf{D}_m[\mathcal{I}_t] d\mathbf{m} \\ &= \text{Ent}_m(\mu) - \int_M \log(\mathbf{D}_m[\mathcal{I}_t]) d\mu. \end{aligned}$$

(To be precise, we applied Theorem 7.3 to $\max\{f, 0\}$ and $\max\{-f, 0\}$ and took their difference.) Hence, by localizing the assumption (7.3), we find that $-\log(\mathbf{D}_m[\mathcal{I}_t](x))$ is K -convex in t for μ -almost every x . Thus the monotone convergence theorem shows that

$$\lim_{t \downarrow 0} \int_M \frac{\log(\mathbf{D}_m[\mathcal{I}_t])}{t} d\mu = \int_M \lim_{t \downarrow 0} \frac{\log(\mathbf{D}_m[\mathcal{I}_t])}{t} d\mu = \int_M \lim_{t \downarrow 0} \frac{\mathbf{D}_m[\mathcal{I}_t] - 1}{t} d\mu.$$

We see by calculation

$$\lim_{t \downarrow 0} \frac{\mathbf{D}_m[\mathcal{T}_t] - 1}{t} = \Delta_m^{\mathbf{H}}\varphi \quad \mu\text{-almost everywhere.} \quad (7.5)$$

Indeed, observe in the calculation in Subsection 4.3 that $d\mathcal{T}_t = B(t)$ and $B(0) = I_n$. Then we have

$$(\det B)'(0) = \text{trace}(B^{-1}\dot{B})(0) = -\text{trace}(B^{-1}A)(0) = \Delta_m^{\mathbf{H}}\varphi.$$

Plugging the effect of the measure \mathbf{m} into this yields (7.5) (recall (4.4)).

Notice that the singular part of $\Delta_m^{\mathbf{H}}\varphi$ is non-negative by looking at the second order term of (4.1). Finally, we obtain (7.4) by integration by parts, see Steps 1–3 in [Vi, Theorem 23.14] for details. \square

For $\mu = \rho\mathbf{m} \in \mathcal{P}(M)$, define

$$I_m(\mu) := \int_M \mathbf{H}(-d[\log \rho]) d\mu \in [0, \infty].$$

This quantity can be thought of as the *Fisher information* adapted to our context.

Theorem 7.6 (Functional inequalities) *Assume that $\mathbf{m}[M] = 1$ and that Ent_m is K -convex in the sense of (7.3) for some $K, T > 0$.*

(i) (Talagrand inequality) *For any $\mu \in \mathcal{P}(M)$, it holds*

$$C_T^{\mathbf{H}}(\mu, \mathbf{m}) \leq \frac{2}{KT} \text{Ent}_m(\mu).$$

(ii) (HWI inequality) *Assume that $\mu = \rho\mathbf{m} \in \mathcal{P}(M)$ with $\rho \in H^1(M; \mathbf{H})$, and let φ be a semi-convex function deriving the optimal transport from μ to \mathbf{m} as in Theorem 7.1. Then we have*

$$\text{Ent}_m(\mu) \leq TI_m(\mu) + T \int_M \mathbf{L}(\nabla\varphi) d\mu - \frac{KT}{2} C_T^{\mathbf{H}}(\mu, \mathbf{m}).$$

Proof. (i) Let $(\mu_t)_{t \in [0, T]}$ be the optimal transport from $\mu_0 = \mu$ to $\mu_T = \mathbf{m}$. It follows from (7.3) and $\text{Ent}_m(\mathbf{m}) = 0$ that, for $t \in [0, 1)$,

$$\text{Ent}_m(\mu_{tT}) \leq (1-t) \text{Ent}_m(\mu) - \frac{K}{2}(1-t)tTC_T^{\mathbf{H}}(\mu, \mathbf{m}). \quad (7.6)$$

Since $\text{Ent}_m(\mu_{tT}) \geq -\log(\mathbf{m}[M]) = 0$, dividing (7.6) by $1-t$ and letting $t \rightarrow 1$ shows the claim.

(ii) We again use (7.6) and find, for $t \in (0, 1]$,

$$\frac{\text{Ent}_m(\mu_{tT}) - \text{Ent}_m(\mu)}{t} \leq -\text{Ent}_m(\mu) - \frac{K}{2}(1-t)TC_T^{\mathbf{H}}(\mu, \mathbf{m}).$$

Proposition 7.5 yields

$$\begin{aligned} \liminf_{t \downarrow 0} \frac{\text{Ent}_{\mathbf{m}}(\mu_{tT}) - \text{Ent}_{\mathbf{m}}(\mu)}{tT} &\geq \int_M \frac{d\rho}{\rho} (\nabla \varphi) d\mu \\ &\geq - \int_M \{ \mathbf{H}(-d[\log \rho]) + \mathbf{L}(\nabla \varphi) \} d\mu \\ &= -I_{\mathbf{m}}(\mu) - \int_M \mathbf{L}(\nabla \varphi) d\mu. \end{aligned}$$

This completes the proof. \square

In the Riemannian or Finsler case, the HWI inequality (with $T = 1$) has the sharper form

$$\text{Ent}_{\mathbf{m}}(\mu) \leq 2\sqrt{I_{\mathbf{m}}(\mu)C_1^{\mathbf{L}}(\mu, \mathbf{m})} - \frac{K}{2}C_1^{\mathbf{H}}(\mu, \mathbf{m})$$

and derives the *log-Sobolev inequality*

$$\text{Ent}_{\mathbf{m}}(\mu) \leq \frac{2}{K}I_{\mathbf{m}}(\mu)$$

by recalling $C_1^{\mathbf{H}} = C_1^{\mathbf{L}}$. See [OT1, §5] and [OT2, §6] for related functional inequalities on different kinds of entropies inspired by information theory.

8 Heat flow

In this section, we study the evolution equation $\partial_t u = \Delta_{\mathbf{m}}^{\mathbf{H}} u$ regarded as the *heat equation* associated with our nonlinear weighted Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}}$. In the Euclidean setting, there are two ways to interpret heat flow as gradient flow. The classical strategy is to consider heat flow as the gradient flow of the Dirichlet energy in the L^2 -space. Another one initiated by the seminal work of Jordan et al. [JKO] is to consider heat flow as the gradient flow of the relative entropy in the L^2 -Wasserstein space. These interpretations and their equivalence were generalized to various settings ([Sa], [Oh2], [GO], [Er], [OS1], [Maa], [GKO], [AGS3] etc.) including singular spaces without differentiable structures.

We shall investigate these two strategies for our heat equation and see that they produce different gradient flows, because of the non-homogeneity of the Hamiltonian \mathbf{H} .

Let M be compact throughout the section, and fix an auxiliary Riemannian metric g of M similarly to the previous section.

8.1 Heat flow as gradient flow of energy

For $(u_t)_{t \geq 0} \subset H^1(M; \mathbf{H}) \cap H^1(M; \mathbf{L})$, we readily see from (2.2) that the energy form $\mathcal{E} := \mathcal{E}_M$ defined in Subsection 4.2 satisfies

$$\partial_{\varepsilon}[\mathcal{E}(u_t + \varepsilon\phi)]|_{\varepsilon=0} = \int_M d\phi(\nabla u_t) d\mathbf{m} = - \int_M \phi \Delta_{\mathbf{m}}^{\mathbf{H}} u_t d\mathbf{m}$$

for every test function $\phi \in \mathcal{C}^{\infty}(M)$ (provided that $\mathcal{E}(u_t + \varepsilon\phi) < \infty$ for some $\varepsilon > 0$). Therefore, with respect to the L^2 -inner product structure, $\Delta_{\mathbf{m}}^{\mathbf{H}} u_t$ is the ‘gradient vector’

of $-\mathcal{E}$ at u_t and the heat equation $\partial_t u = \Delta_{\mathbf{m}}^{\mathbf{H}} u$ coincides with the (descending) gradient flow equation of the potential function \mathcal{E} .

Motivated by the above heuristic argument, we shall build weak solutions to $\partial_t u = \Delta_{\mathbf{m}}^{\mathbf{H}} u$ as gradient curves of \mathcal{E} . In fact, since \mathcal{E} is a convex and lower semi-continuous functional on $L^2(M; \mathbf{m})$ (Lemma 4.2), the classical theory due to Brézis, Crandall, Liggett et al applies (see [Bré], [CL]). For completeness, we repeat the construction in [OS1, §3] concerning the Finsler case with the help of [Oh2]. Because of the author's familiarity, our construction follows the line of [May] which deals with the more general situation of convex functions on nonpositively curved metric spaces.

For $u \in H^1(M; \mathbf{H})$, define

$$|\nabla(-\mathcal{E})|(u) := \max \left\{ \limsup_{\bar{u} \rightarrow u} \frac{\mathcal{E}(u) - \mathcal{E}(\bar{u})}{\|\bar{u} - u\|_{L^2}}, 0 \right\},$$

where $\bar{u} \in H^1(M; \mathbf{H})$ and $\bar{u} \rightarrow u$ is in $L^2(M; \mathbf{m})$. By the convexity of \mathcal{E} , we have $|\nabla(-\mathcal{E})|(u) = 0$ if and only if $\mathcal{E}(u) = 0$, i.e., u is constant. Similarly to [OS1, Lemma 3.3], we can find a unique direction attaining $|\nabla(-\mathcal{E})|(u)$.

Lemma 8.1 *If $0 < |\nabla(-\mathcal{E})|(u) < \infty$, then there exists a unique element $w \in L^2(M; \mathbf{m})$ satisfying $\|w\|_{L^2} = |\nabla(-\mathcal{E})|(u)$ and*

$$\limsup_{\bar{w} \rightarrow w} \lim_{t \downarrow 0} \frac{\mathcal{E}(u) - \mathcal{E}(u + t\bar{w})}{t\|\bar{w}\|_{L^2}} = |\nabla(-\mathcal{E})|(u), \quad (8.1)$$

where the convergence $\bar{w} \rightarrow w$ is in $L^2(M; \mathbf{m})$.

Proof. Take a sequence $\{\hat{w}_i\}_{i \in \mathbb{N}} \subset L^2(M; \mathbf{m}) \setminus \{0\}$ such that $\hat{w}_i \rightarrow 0$ and

$$\lim_{i \rightarrow \infty} \frac{\mathcal{E}(u) - \mathcal{E}(u + \hat{w}_i)}{\|\hat{w}_i\|_{L^2}} = |\nabla(-\mathcal{E})|(u),$$

and normalize it as

$$w_i := \frac{|\nabla(-\mathcal{E})|(u)}{\|\hat{w}_i\|_{L^2}} \hat{w}_i.$$

It follows from the convexity of \mathcal{E} that

$$\liminf_{i \rightarrow \infty} \lim_{t \downarrow 0} \frac{\mathcal{E}(u) - \mathcal{E}(u + tw_i)}{t\|w_i\|_{L^2}} \geq \lim_{i \rightarrow \infty} \frac{\mathcal{E}(u) - \mathcal{E}(u + \hat{w}_i)}{\|\hat{w}_i\|_{L^2}} = |\nabla(-\mathcal{E})|(u).$$

Thus we have, together with the definition of $|\nabla(-\mathcal{E})|(u)$,

$$\lim_{i \rightarrow \infty} \lim_{t \downarrow 0} \frac{\mathcal{E}(u) - \mathcal{E}(u + tw_i)}{t\|w_i\|_{L^2}} = |\nabla(-\mathcal{E})|(u). \quad (8.2)$$

The convexity of \mathcal{E} also implies that, for any $i, j \geq 1$,

$$\begin{aligned} & \lim_{t \downarrow 0} \frac{\mathcal{E}(u) - \mathcal{E}(u + t(w_i + w_j)/2)}{t\|(w_i + w_j)/2\|_{L^2}} \\ & \geq \frac{|\nabla(-\mathcal{E})|(u)}{\|(w_i + w_j)/2\|_{L^2}} \lim_{t \downarrow 0} \left\{ \frac{\mathcal{E}(u) - \mathcal{E}(u + tw_i)}{2t\|w_i\|_{L^2}} + \frac{\mathcal{E}(u) - \mathcal{E}(u + tw_j)}{2t\|w_j\|_{L^2}} \right\}, \end{aligned}$$

we used $\|w_i\|_{L^2} = \|w_j\|_{L^2} = |\nabla(-\mathcal{E})|(u)$. Since the LHS is not greater than $|\nabla(-\mathcal{E})|(u)$, we find from (8.2) that $\liminf_{i,j \rightarrow \infty} \|w_i + w_j\|_{L^2} \geq 2|\nabla(-\mathcal{E})|(u)$. Combining this with $\|w_i\|_{L^2} = |\nabla(-\mathcal{E})|(u)$ for all i implies that $\{w_i\}_{i \in \mathbb{N}}$ is a Cauchy sequence in $L^2(M; \mathbf{m})$ and converges to some element $w \in L^2(M; \mathbf{m})$. The desired equation (8.1) holds due to (8.2).

The uniqueness follows from the strict convexity of $\|\cdot\|_{L^2}$ and the convexity of \mathcal{E} . \square

We excluded the case of $|\nabla(-\mathcal{E})|(u) = 0$ since (8.1) does not necessarily hold with $w = 0$.

Definition 8.2 (Gradient vectors of $-\mathcal{E}$) At $u \in H^1(M; \mathbf{H})$ with $0 < |\nabla(-\mathcal{E})|(u) < \infty$, we define the *gradient vector* of $-\mathcal{E}$ by $\nabla(-\mathcal{E})(u) := w \in L^2(M; \mathbf{m})$, where w is given in Lemma 8.1. We also set $\nabla(-\mathcal{E})(u) := 0 \in L^2(M; \mathbf{m})$ if $|\nabla(-\mathcal{E})|(u) = 0$.

A gradient curve of \mathcal{E} should be understood as a solution to $\partial_t u = \nabla(-\mathcal{E})(u)$. We explain how to formulate and construct it. Fix $u_0 \in H^1(M; \mathbf{H})$ and $\delta > 0$. Denote by $U_\delta(u_0) \in H^1(M; \mathbf{H})$ the unique minimizer of the strictly convex, lower semi-continuous functional

$$u \mapsto \mathcal{E}(u) + \frac{\|u - u_0\|_{L^2}^2}{2\delta}. \quad (8.3)$$

Then, as $k \rightarrow \infty$, the discrete approximation scheme $(U_{t/k})^k(u_0)$ converges to a curve $(u_t)_{t>0} \subset L^2(M; \mathbf{m})$ such that $\lim_{t \downarrow 0} u_t = u_0$ in $L^2(M; \mathbf{m})$ and $\mathcal{E}(u_t) \leq \mathcal{E}(u_0)$ for all $t > 0$ (we refer to [May, Theorem 1.13] for details). Moreover, the convergence is uniform on $(0, T]$ for each $T > 0$, and $(u_t)_{t \geq 0}$ satisfies the following:

- The curve $(u_t)_{t>0}$ is locally Lipschitz in $L^2(M; \mathbf{m})$ and satisfies

$$\lim_{\delta \downarrow 0} \frac{\|u_{t+\delta} - u_t\|_{L^2}}{\delta} = |\nabla(-\mathcal{E})|(u_t) \quad (8.4)$$

at all $t \geq 0$ ([May, Theorems 2.9, 2.17]). In particular, $|\nabla(-\mathcal{E})|(u_t) < \infty$ holds and the gradient vector $\nabla(-\mathcal{E})(u_t) \in L^2(M; \mathbf{m})$ is well-defined for all $t > 0$.

- The function $t \mapsto \mathcal{E}(u_t)$ is continuous on $[0, \infty)$ and locally Lipschitz on $(0, \infty)$ ([May, Corollary 2.10]). Combining (8.4) with [May, Lemma 2.11, Theorem 2.14], we have, for all $t \geq 0$,

$$\lim_{\delta \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(u_{t+\delta})}{\delta} = |\nabla(-\mathcal{E})|(u_t)^2. \quad (8.5)$$

With the help of (8.4) and (8.5), the same discussion as Lemma 8.1 (replacing \hat{w}_i and w_i with $u_{t+\delta} - u_t$ and $(u_{t+\delta} - u_t)/\delta$, respectively) shows that

$$\lim_{\delta \downarrow 0} \left\| \frac{u_{t+\delta} - u_t}{\delta} - \nabla(-\mathcal{E})(u_t) \right\|_{L^2} = 0 \quad (8.6)$$

for all $t > 0$, i.e., $\partial_t u_t = \nabla(-\mathcal{E})(u_t)$ in the weak sense. In addition, we obtain the following similarly to [Oh2, Lemma 6.4].

Lemma 8.3 (i) Fix arbitrary $t > 0$ and put $\sigma_\delta := \mathbf{U}_\delta(u_t)$ by suppressing the dependence on t . Then we have

$$\lim_{\delta \downarrow 0} \frac{\|u_{t+\delta} - \sigma_\delta\|_{L^2}}{\delta} = 0.$$

(ii) Moreover, $d\sigma_\delta$ converges to du_t as $\delta \downarrow 0$ in the L^2 -norm with respect to g .

Proof. (i) Recall that $|\nabla(-\mathcal{E})|(u_t) < \infty$. If $|\nabla(-\mathcal{E})|(u_t) = 0$, then $\mathcal{E}(u_t) = 0$ and hence $\sigma_\delta = u_t$ for all $\delta > 0$. Thus the claim follows from (8.4).

Suppose $|\nabla(-\mathcal{E})|(u_t) > 0$. By (8.6), the claim is equivalent to

$$\lim_{\delta \downarrow 0} \left\| \frac{\sigma_\delta - u_t}{\delta} - \nabla(-\mathcal{E})(u_t) \right\|_{L^2} = 0.$$

Take a sequence $\{w_i\}_{i \in \mathbb{N}} \subset L^2(M; \mathfrak{m})$ as in the proof of Lemma 8.1, namely $\|w_i\|_{L^2} = |\nabla(-\mathcal{E})|(u_t)$ for all i and

$$\lim_{i \rightarrow \infty} \lim_{s \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(u_t + sw_i)}{s\|w_i\|_{L^2}} = |\nabla(-\mathcal{E})|(u_t).$$

Put

$$w_i^\delta := u_t + \frac{\|\sigma_\delta - u_t\|_{L^2}}{\|w_i\|_{L^2}} w_i$$

and observe from the choice of σ_δ that

$$\liminf_{\delta \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(\sigma_\delta)}{\|\sigma_\delta - u_t\|_{L^2}} \geq \lim_{\delta \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(w_i^\delta)}{\|\sigma_\delta - u_t\|_{L^2}} = \lim_{s \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(u_t + sw_i)}{s\|w_i\|_{L^2}}.$$

Letting $i \rightarrow \infty$ implies

$$\liminf_{\delta \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(\sigma_\delta)}{\|\sigma_\delta - u_t\|_{L^2}} \geq |\nabla(-\mathcal{E})|(u_t),$$

and hence

$$\lim_{\delta \downarrow 0} \frac{\mathcal{E}(u_t) - \mathcal{E}(\sigma_\delta)}{\|\sigma_\delta - u_t\|_{L^2}} = |\nabla(-\mathcal{E})|(u_t). \quad (8.7)$$

If we have $\lim_{\delta \downarrow 0} \|\sigma_\delta - u_t\|_{L^2}/\delta = |\nabla(-\mathcal{E})|(u_t)$, then we can complete the proof along the same line as Lemma 8.1. To see this, we observe

$$\mathcal{E}(\sigma_\delta) + \frac{\|\sigma_\delta - u_t\|_{L^2}^2}{2\delta} \leq \mathcal{E}(u_t + \delta w_i) + \frac{\delta^2 |\nabla(-\mathcal{E})|(u_t)^2}{2\delta}$$

by the choice of σ_δ , and

$$\mathcal{E}(\sigma_\delta) \geq \mathcal{E}(u_t) - |\nabla(-\mathcal{E})|(u_t) \|\sigma_\delta - u_t\|_{L^2}$$

by the convexity of \mathcal{E} . Thus we find

$$\begin{aligned} \frac{\|\sigma_\delta - u_t\|_{L^2}^2}{2\delta^2} &\leq \frac{\mathcal{E}(u_t + \delta w_i) - \mathcal{E}(\sigma_\delta)}{\delta} + \frac{|\nabla(-\mathcal{E})|(u_t)^2}{2} \\ &\leq \frac{\mathcal{E}(u_t + \delta w_i) - \mathcal{E}(u_t)}{\delta} + \frac{\|\sigma_\delta - u_t\|_{L^2}}{\delta} |\nabla(-\mathcal{E})|(u_t) + \frac{|\nabla(-\mathcal{E})|(u_t)^2}{2}. \end{aligned}$$

Letting $\delta \downarrow 0$ and then $i \rightarrow \infty$, we obtain

$$\left(\limsup_{\delta \downarrow 0} \frac{\|\sigma_\delta - u_t\|_{L^2}}{\delta} \right)^2 \leq 2|\nabla(-\mathcal{E})|(u_t) \cdot \limsup_{\delta \downarrow 0} \frac{\|\sigma_\delta - u_t\|_{L^2}}{\delta} - |\nabla(-\mathcal{E})|(u_t)^2.$$

Therefore $\limsup_{\delta \downarrow 0} \|\sigma_\delta - u_t\|_{L^2}/\delta = |\nabla(-\mathcal{E})|(u_t)$ and similarly $\liminf_{\delta \downarrow 0} \|\sigma_\delta - u_t\|_{L^2}/\delta = |\nabla(-\mathcal{E})|(u_t)$.

(ii) The strict convexity of \mathbf{H} and (8.7) show that $\{d\sigma_\delta\}_{\delta>0}$ is a Cauchy sequence and converges to du_t , i.e., $\lim_{\delta \downarrow 0} \int_M |d\sigma_\delta - du_t|_g^2 d\mathbf{m} = 0$. (We remark that, however, $\int_M |d\sigma_\delta|_g^2 d\mathbf{m} = \infty$ may happen.) \square

We are ready to show the main result of the subsection. Due to technical difficulties, we impose the following assumption:

(A) For each $\varepsilon > 0$, there is $C_\varepsilon > 0$ such that $|\tau^*(\alpha) - \tau^*(\beta)|_g \leq C_\varepsilon|\alpha - \beta|_g$ for all $\alpha, \beta \in T_x^*M$, $x \in M$, with $|\tau^*(\alpha) - \tau^*(\beta)|_g \geq \varepsilon$.

This is obviously a strong constraint since it is global in the vertical direction, however, we need this kind of condition to control the behavior of $\nabla\sigma_\delta$ by that of $d\sigma_\delta$.

Theorem 8.4 (Heat flow as gradient flow of energy) *Suppose (A). Given an initial function $u_0 \in H^1(M; \mathbf{H})$, let $(u_t)_{t \geq 0} \subset H^1(M; \mathbf{H})$ be a gradient curve of \mathcal{E} constructed as above. Then $(u_t)_{t \geq 0}$ solves the heat equation $\partial_t u_t = \Delta_{\mathbf{m}}^{\mathbf{H}} u_t$ in the weak sense that*

$$\int_0^\infty \int_M \partial_t \phi \cdot u_t d\mathbf{m} dt = \int_0^\infty \int_M d\phi_t(\nabla u_t) d\mathbf{m} dt \quad (8.8)$$

for all $\phi \in \mathcal{C}_c^\infty((0, \infty) \times M)$, where $\phi_t := \phi(t, \cdot)$.

Moreover, for all $t > 0$, the distributional Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}} u_t$ is absolutely continuous with respect to \mathbf{m} and its density function coincides with $\nabla(-\mathcal{E})(u_t)$. In particular, we have

$$\|\Delta_{\mathbf{m}}^{\mathbf{H}} u_t\|_{L^2} = \|\nabla(-\mathcal{E})(u_t)\|_{L^2} = |\nabla(-\mathcal{E})|(u_t)$$

for all $t > 0$.

Proof. We fix a test function $\phi \in \mathcal{C}_c^\infty((0, \infty) \times M)$ and shall show that

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \int_M (\phi_{t+\delta} u_{t+\delta} - \phi_t u_t) d\mathbf{m} = \int_M \{\partial_t \phi_t \cdot u_t - d\phi_t(\nabla u_t)\} d\mathbf{m} \quad (8.9)$$

for all $t > 0$, which implies (8.8) by integration. We remark that the RHS is well-defined because

$$|d\phi_t(\nabla u_t)| \leq |d\phi_t|_g |\nabla u_t|_g \leq |d\phi_t|_g (\varepsilon + C_\varepsilon |du_t|_g)$$

by (A) and

$$\int_M |du_t|_g d\mathbf{m} \leq \int_M \{C_1 \mathbf{H}(du_t) + C_2\} d\mathbf{m} < \infty$$

for some $C_1, C_2 > 0$ (super-linearity).

Fix $t > 0$. For small $\delta, \varepsilon > 0$, let us compare $\sigma_\delta := \mathbf{U}_\delta(u_t)$ and $\sigma_\delta^\varepsilon := \sigma_\delta + \varepsilon\phi_t$. By the definition of $\mathbf{U}_\delta(u_t)$, we have

$$\mathcal{E}(\sigma_\delta) + \frac{\|\sigma_\delta - u_t\|_{L^2}^2}{2\delta} \leq \mathcal{E}(\sigma_\delta^\varepsilon) + \frac{\|\sigma_\delta^\varepsilon - u_t\|_{L^2}^2}{2\delta}.$$

On the one hand, by expanding $\|\sigma_\delta - u_t\|_{L^2}^2$ and $\|\sigma_\delta^\varepsilon - u_t\|_{L^2}^2$, we find

$$\lim_{\varepsilon \downarrow 0} \frac{\|\sigma_\delta^\varepsilon - u_t\|_{L^2}^2 - \|\sigma_\delta - u_t\|_{L^2}^2}{2\varepsilon} = \int_M \phi_t(\sigma_\delta - u_t) d\mathbf{m}.$$

On the other hand, we deduce from the convexity of \mathbf{H} that

$$\lim_{\varepsilon \downarrow 0} \frac{\mathcal{E}(\sigma_\delta^\varepsilon) - \mathcal{E}(\sigma_\delta)}{\varepsilon} = \int_M \lim_{\varepsilon \downarrow 0} \frac{\mathbf{H}(d\sigma_\delta^\varepsilon) - \mathbf{H}(d\sigma_\delta)}{\varepsilon} d\mathbf{m} = \int_M d\phi_t(\nabla\sigma_\delta) d\mathbf{m}.$$

Therefore we have, with the help of Lemma 8.3(ii) and **(A)**,

$$\liminf_{\delta \downarrow 0} \frac{1}{\delta} \int_M \phi_t(\sigma_\delta - u_t) d\mathbf{m} \geq -\lim_{\delta \downarrow 0} \int_M d\phi_t(\nabla\sigma_\delta) d\mathbf{m} = -\int_M d\phi_t(\nabla u_t) d\mathbf{m}.$$

To be precise, for each $\varepsilon > 0$, we have by **(A)**

$$\left| \int_M d\phi_t(\nabla\sigma_\delta - \nabla u_t) d\mathbf{m} \right| \leq \int_M |d\phi_t|_g(\varepsilon + C_\varepsilon|d\sigma_\delta - du_t|_g) d\mathbf{m}$$

which tends to 0 as $\delta \downarrow 0$ and then $\varepsilon \downarrow 0$. Together with Lemma 8.3(i), we obtain

$$\begin{aligned} \liminf_{\delta \downarrow 0} \frac{1}{\delta} \int_M (\phi_{t+\delta}u_{t+\delta} - \phi_t u_t) d\mathbf{m} &= \liminf_{\delta \downarrow 0} \frac{1}{\delta} \int_M \{(\phi_{t+\delta} - \phi_t)u_{t+\delta} + \phi_t(u_{t+\delta} - u_t)\} d\mathbf{m} \\ &= \int_M \partial_t \phi_t \cdot u_t d\mathbf{m} + \liminf_{\delta \downarrow 0} \frac{1}{\delta} \int_M \phi_t(\sigma_\delta - u_t) d\mathbf{m} \\ &\geq \int_M \{\partial_t \phi_t \cdot u_t - d\phi_t(\nabla u_t)\} d\mathbf{m}. \end{aligned}$$

By exchanging ϕ with $-\phi$, we obtain the desired equation (8.9).

For a time-independent test function $\phi \in \mathcal{C}^\infty(M)$, (8.6) and (8.9) show that

$$\int_M \phi \nabla(-\mathcal{E})(u_t) d\mathbf{m} = \lim_{\delta \downarrow 0} \frac{1}{\delta} \int_M \phi(u_{t+\delta} - u_t) d\mathbf{m} = -\int_M d\phi(\nabla u_t) d\mathbf{m}$$

for all $t > 0$. Therefore the distributional Laplacian $\Delta_{\mathbf{m}}^{\mathbf{H}} u_t$ is absolutely continuous with respect to \mathbf{m} and the density function is $\nabla(-\mathcal{E})(u_t)$. \square

Given two solutions $(u_t)_{t \geq 0}$, $(\bar{u}_t)_{t \geq 0}$ to the heat equation constructed in Theorem 8.4, applying (8.9) with $\phi = u - \bar{u}$ (this is possible by [OS1, Remark 3.2(i)]), we find

$$\begin{aligned} \frac{d}{dt} [\|u_t - \bar{u}_t\|_{L^2}^2] &= \frac{d}{dt} \left[\int_M \{(u_t - \bar{u}_t)u_t - (u_t - \bar{u}_t)\bar{u}_t\} d\mathbf{m} \right] \\ &= \int_M \{\partial_t(u_t - \bar{u}_t)(u_t - \bar{u}_t) - d(u_t - \bar{u}_t)(\nabla u_t - \nabla \bar{u}_t)\} d\mathbf{m} \\ &= -2 \int_M (du_t - d\bar{u}_t)(\nabla u_t - \nabla \bar{u}_t) d\mathbf{m} \end{aligned}$$

at almost every $t > 0$. The convexity of \mathbf{H} implies

$$(du_t - d\bar{u}_t)(\nabla u_t - \nabla \bar{u}_t) = (du_t - d\bar{u}_t) \left(\sum_{i=1}^n \mathbf{H}_{\alpha_i}(du_t) \partial_{x^i} - \sum_{i=1}^n \mathbf{H}_{\alpha_i}(d\bar{u}_t) \partial_{x^i} \right) \geq 0,$$

which yields the contraction $(\|u_t - \bar{u}_t\|_{L^2}^2)' \leq 0$ of the heat flow. In particular, if $u_0 = \bar{u}_0$ \mathbf{m} -almost everywhere, then $u_t = \bar{u}_t$ \mathbf{m} -almost everywhere for all $t > 0$.

8.2 Gradient flow of the relative entropy

We next consider the gradient flow of the relative entropy $\text{Ent}_{\mathbf{m}}$ with respect to the Lagrangian cost function $C_T^{\mathbf{L}}$ introduced in Section 7. We will verify that such a flow is produced from weak solutions to the evolution equation

$$\partial_t u = -\text{div}_{\mathbf{m}}(u \nabla[-\log u]), \quad (8.10)$$

which is different from the heat equation $\partial_t u = \Delta_{\mathbf{m}}^{\mathbf{H}} u$ due to the non-homogeneity of the gradient operator ∇ .

Let us begin with a discussion on how to define gradient flows. Our strategy follows the metric approach recently intensively studied by Ambrosio, Gigli, Savaré and others (see [AGS1]). Given a C^1 -function $f : M \rightarrow \mathbb{R}$, the *gradient flow* of f with respect to \mathbf{L} (or \mathbf{H}) is introduced as the family of maps $\{\mathcal{G}_t\}_{t \geq 0}$ sending x to $\mathcal{G}_t(x) = \eta(t)$, where η is the C^1 -curve solving

$$\dot{\eta}(t) = \nabla(-f)(\eta(t)), \quad \eta(0) = x. \quad (8.11)$$

Since $v = \nabla(-f)(\eta(t))$ is the unique element satisfying

$$-df_{\eta(t)}(v) = \mathbf{H}(-df_{\eta(t)}) + \mathbf{L}(v)$$

in $T_{\eta(t)}M$ and $-df_{\eta(t)}(v) \leq \mathbf{H}(-df_{\eta(t)}) + \mathbf{L}(v)$ holds in general, the gradient flow equation (8.11) can be rewritten as

$$(f \circ \eta)'(t) \leq -\mathbf{H}(-df_{\eta(t)}) - \mathbf{L}(\dot{\eta}(t)).$$

Integrating this inequality leads to

$$f(\eta(t)) \leq f(\eta(s)) - \int_s^t \{ \mathbf{H}(-df_{\eta(r)}) + \mathbf{L}(\dot{\eta}(r)) \} dr$$

for all $0 \leq s < t$. In order to avoid the use of the derivative df , we observe that, at any $y \in M$,

$$\mathbf{H}(-df_y) = \sup_{v \in T_y M} \{ -df_y(v) - \mathbf{L}(v) \} = \sup_{\gamma} \left\{ \lim_{t \downarrow 0} \frac{f(y) - f(\gamma(t))}{t} - \mathbf{L}(\dot{\gamma}(0)) \right\},$$

where γ runs over all differentiable curves with $\gamma(0) = y$. Note finally that

$$\mathbf{L}(\dot{\eta}(r)) = \lim_{\delta \downarrow 0} \frac{c_{\delta}^{\mathbf{L}}(\eta(r), \eta(r + \delta))}{\delta}.$$

Then, it is natural to introduce the following metric definition.

Definition 8.5 (Gradient flow of Ent_m) A curve $(\mu_t)_{t \geq 0} \subset \mathcal{P}(M)$ which is continuous in the weak topology of $\mathcal{P}(M)$ is called a *gradient curve* of Ent_m if $\text{Ent}_m(\mu_t) < \infty$ for all $t > 0$ and if

$$\text{Ent}_m(\mu_t) \leq \text{Ent}_m(\mu_s) - \int_s^t \{ |d(-\text{Ent}_m)|_{\mathbf{H}}(\mu_r) + |\dot{\mu}_r|_{\mathbf{L}} \} dr \quad (8.12)$$

holds for all $0 \leq s < t$. Here we set

$$|\dot{\mu}_r|_{\mathbf{L}} := \limsup_{\delta \downarrow 0} \frac{C_\delta^{\mathbf{L}}(\mu_r, \mu_{r+\delta})}{\delta},$$

$$|d(-\text{Ent}_m)|_{\mathbf{H}}(\nu) := \sup_{(\nu_s)_{s \in [0, T]}} \left\{ \limsup_{s \downarrow 0} \frac{\text{Ent}_m(\nu) - \text{Ent}_m(\nu_s)}{s} - |\dot{\nu}_0|_{\mathbf{L}} \right\}$$

for $\nu \in \mathcal{P}(M)$ with $\text{Ent}_m(\nu) < \infty$, where $(\nu_s)_{s \in [0, T]}$ runs over all optimal transports with $\nu_0 = \nu$ (possibly $\nu_t \equiv \nu$, so that $|d(-\text{Ent}_m)|_{\mathbf{H}} \geq 0$).

We assume that Ent_m is K -convex in the sense of (7.3) for some $K \in \mathbb{R}$ and all $T > 0$, and shall show that, for a weak solution $(\rho_t)_{t \geq 0}$ to (8.10) with $\mu_t := \rho_t \mathbf{m} \in \mathcal{P}(M)$, the curve $(\mu_t)_{t \geq 0}$ enjoys (8.12). For simplicity, suppose that the following condition is satisfied:

- (C) ρ_t is locally g -Lipschitz on $(0, \infty) \times M$, $\lim_{r \downarrow 0} d\rho_{t+r} = d\rho_t$ in the L^2 -norm with respect to g and $\inf_M \rho_t > 0$ for every $t > 0$ (then clearly $\text{Ent}_m(\mu_t) < \infty$ for $t > 0$).

Observe that the heat flow constructed in the previous subsection satisfies $\lim_{r \downarrow 0} d\rho_{t+r} = d\rho_t$ similarly to Lemma 8.3(ii).

First of all, it follows from Proposition 7.5 that

$$|d(-\text{Ent}_m)|_{\mathbf{H}}(\mu) \leq \int_M \mathbf{H} \left(-\frac{d\rho}{\rho} \right) d\mu. \quad (8.13)$$

Next, observe from (8.10) (with the test function $\log \rho_t$) that

$$\frac{d}{dt} [\text{Ent}_m(\mu_t)] = \int_M d(\log \rho_t)(\rho_t \nabla[-\log \rho_t]) d\mathbf{m} = \int_M \frac{d\rho_t}{\rho_t} (\nabla[-\log \rho_t]) d\mu_t. \quad (8.14)$$

Finally, $|\dot{\mu}_t|_{\mathbf{L}}$ is estimated similarly to [GKO, Proposition 3.7] as follows.

Lemma 8.6 *At each $t > 0$, it holds that*

$$|\dot{\mu}_t|_{\mathbf{L}} \leq \int_M \mathbf{L}(\nabla[-\log \rho_t]) d\mu_t.$$

Proof. For $s > 0$, we have by [BB, Proposition 21]

$$C_s^{\mathbf{L}}(\mu_t, \mu_{t+s}) = \sup_{\varphi} \left\{ \int_M \varphi_{t+s} d\mu_{t+s} - \int_M \varphi_t d\mu_t \right\},$$

where the supremum is taken over all g -Lipschitz functions $\varphi : [t, t + s] \times M \longrightarrow \mathbb{R}$ which are \mathcal{C}^1 on $(t, t + s) \times M$ and satisfy the inequality (of Hamilton–Jacobi type)

$$\partial_r \varphi + \mathbf{H}(d\varphi) \leq 0 \quad \text{on } (t, t + s) \times M.$$

Together with (8.10), we obtain

$$\begin{aligned} \int_M \varphi_{t+s} d\mu_{t+s} - \int_M \varphi_t d\mu_t &= \int_M \int_0^s \partial_r [\varphi_{t+r} \rho_{t+r}] dr d\mathbf{m} \\ &\leq \int_0^s \int_M \{-\mathbf{H}(d\varphi_{t+r}) + d\varphi_{t+r}(\nabla[-\log \rho_{t+r}])\} d\mu_{t+r} dr \\ &\leq \int_0^s \int_M \mathbf{L}(\nabla[-\log \rho_{t+r}]) d\mu_{t+r} dr. \end{aligned}$$

Therefore we conclude that

$$|\dot{\mu}_t|_{\mathbf{L}} \leq \lim_{s \downarrow 0} \frac{1}{s} \int_0^s \int_M \mathbf{L}(\nabla[-\log \rho_{t+r}]) d\mu_{t+r} dr = \int_M \mathbf{L}(\nabla[-\log \rho_t]) d\mu_t$$

by the bounded convergence theorem (since ρ_t is locally g -Lipschitz and $\inf_M \rho_t > 0$) and $\lim_{r \downarrow 0} d\rho_{t+r} = d\rho_t$. \square

Combining above three estimates, we immediately obtain the following.

Theorem 8.7 (Gradient flow of $\text{Ent}_{\mathbf{m}}$) *Assume that $\text{Ent}_{\mathbf{m}}$ is K -convex in the sense of (7.3) for some $K \in \mathbb{R}$ and all $T > 0$. Let $(\rho_t)_{t \geq 0}$ be a weak solution to the equation (8.10) satisfying (C) and $\mu_t := \rho_t \mathbf{m} \in \mathcal{P}(M)$ for all $t \geq 0$. Then $(\mu_t)_{t \geq 0}$ is a gradient curve of $\text{Ent}_{\mathbf{m}}$ in the sense of (8.12).*

Proof. Since $d\rho_t/\rho_t = d[\log \rho_t]$, we deduce from (8.14), (8.13) and Lemma 8.6 that

$$\begin{aligned} \frac{d}{dt} [\text{Ent}_{\mathbf{m}}(\mu_t)] &= - \int_M \{\mathbf{H}(-d[\log \rho_t]) + \mathbf{L}(\nabla[-\log \rho_t])\} d\mu_t \\ &\leq -|d(-\text{Ent}_{\mathbf{m}})|_{\mathbf{H}}(\mu_t) - |\dot{\mu}_t|_{\mathbf{L}}. \end{aligned}$$

This completes the proof. \square

Example 8.8 For the p -homogeneous deformation $\mathbf{H} = |\cdot|_g^p/p$ of a Riemannian metric g , the equation (8.10) turns out

$$\partial_t u = -\text{div}_{\mathbf{m}}(u |\nabla^g[-\log u]|_g^{p-2} \cdot \nabla^g[-\log u]) = \text{div}_{\mathbf{m}} \left(\left| \frac{\nabla^g u}{u} \right|_g^{p-2} \cdot \nabla^g u \right).$$

This recovers the case of $\mathbf{L}(v) = |v|^q/q$ on \mathbb{R}^n studied in [AGS1, §11.3] and [Agu], in which more general entropy functionals are investigated. For instance, employing the *Rényi–Tsallis entropy*

$$S_k(\rho \mathbf{m}) := \frac{1}{k(k-1)} \int_M \rho^k d\mathbf{m}, \quad k = \frac{2p-3}{p-1} = 3-q$$

leads to the p -Laplacian equation: $\partial_t u = \Delta_{\mathbf{m}}^{\mathbf{H}} u = \operatorname{div}_{\mathbf{m}}(|\nabla^g u|_g^{p-2} \cdot \nabla^g u)$. Indeed, Proposition 7.5 is read as

$$\liminf_{t \downarrow 0} \frac{S_k(\mu_t) - S_k(\mu)}{t} \geq \int_M \rho^{k-2} d\rho(\nabla \varphi) d\mu,$$

and then (8.13), (8.14) and Lemma 8.6 become

$$\begin{aligned} |d(-S_k)|_{\mathbf{H}}(\mu) &\leq \int_M \mathbf{H}(-\rho^{k-2} d\rho) d\mu = \frac{1}{p} \int_M |d\rho|_g^p \rho^{(k-2)p} d\mu, \\ \frac{d}{dt}[S_k(\mu_t)] &= - \int_M d\rho_t(\nabla \rho_t) \rho_t^{k-2} d\mathbf{m} = - \int_M \{\mathbf{H}(d\rho_t) + \mathbf{L}(\nabla \rho_t)\} \rho_t^{k-2} d\mathbf{m}, \\ |\dot{\mu}_t|_{\mathbf{L}} &\leq \int_M \mathbf{L}\left(-\frac{\nabla \rho_t}{\rho_t}\right) d\mu_t = \frac{1}{q} \int_M |\nabla^g \rho_t|_g^p \rho_t^{-q} d\mu_t \end{aligned}$$

(compare the last inequality with [AGS2, Lemma 7.2]). Since $(k-2)p = -q$, we find as desired

$$\frac{d}{dt}[S_k(\mu_t)] = - \int_M \left(\frac{|d\rho_t|_g^p}{p} + \frac{|\nabla^g \rho_t|_g^p}{q} \right) \rho_t^{-q} d\mu_t \leq -|d(-S_k)|_{\mathbf{H}}(\mu_t) - |\dot{\mu}_t|_{\mathbf{L}}.$$

In the Finsler case, the equation (8.10) coincides with the heat equation with respect to the reverse Finsler structure $\overleftarrow{F}(v) := F(-v)$, $\partial_t \rho = -\Delta_{\mathbf{m}}^F(-\rho)$, as shown in [OS1]. In general, it is unclear if one can regard (8.10) as the heat equation of some Hamiltonian.

A Appendix

In this appendix, for self-containedness, we give a way of constructing canonical frames as well as a detailed calculation of the curvature operator in coordinates. In both subjects, we follow the line of [Le2] with proofs.

A.1 How to construct canonical frames

Fix $\alpha \in T_x^* M \setminus \{0\}$ and put $\boldsymbol{\alpha}(t) := \Phi_t(\alpha)$ for $t \in (-\varepsilon, \varepsilon)$ throughout the appendix. Our starting ingredient is a family of smooth curves $\boldsymbol{\xi}_i(t) = \sum_{j=1}^n \xi_i^j(t) \partial_{\alpha_j} \in \mathcal{V}_{\boldsymbol{\alpha}(t)}$, $i = 1, \dots, n$, orthonormal with respect to the inner product introduced in Lemma 3.2, namely

$$\sum_{k,l=1}^n \mathbf{H}_{\alpha_k \alpha_l}(\boldsymbol{\alpha}(t)) \xi_i^k(t) \xi_j^l(t) = \delta_{ij} \quad \text{for all } i, j \text{ and } t.$$

We first show that the curves $\bar{e}_i^t := (d\Phi_t)^{-1}(\boldsymbol{\xi}_i(t)) \in \mathcal{J}_{\alpha}^t$ satisfy the first condition in (3.1).

Lemma A.1 *We have $\omega(\bar{e}_i^t, \bar{e}_j^t) = \delta_{ij}$ for all i, j and t .*

Proof. Note that, by $\dot{\Phi}_t = \vec{\mathbf{H}} \circ \Phi_t$,

$$\begin{aligned} \dot{\bar{e}}_i^t &= (d\Phi_t)^{-1} \left(\sum_{k=1}^n \dot{\xi}_i^k(t) \partial_{\alpha_k} \right) \\ &\quad - (d\Phi_t)^{-1} \left[\sum_{k,l=1}^n \xi_i^l(t) \{ \mathbf{H}_{\alpha_k \alpha_l}(\boldsymbol{\alpha}(t)) \partial_{x^k} - \mathbf{H}_{x^k \alpha_l}(\boldsymbol{\alpha}(t)) \partial_{\alpha_k} \} \right]. \end{aligned}$$

Therefore we have, by (2.3),

$$\begin{aligned} \omega(\dot{\bar{e}}_i^t, \bar{e}_j^t) &= -\omega \left(\sum_{k,l=1}^n \mathbf{H}_{\alpha_k \alpha_l}(\boldsymbol{\alpha}(t)) \xi_i^l(t) \partial_{x^k}, \boldsymbol{\xi}_j(t) \right) = \sum_{k,l=1}^n \mathbf{H}_{\alpha_k \alpha_l}(\boldsymbol{\alpha}(t)) \xi_i^l(t) \xi_j^k(t) \\ &= \delta_{ij}. \end{aligned}$$

□

One can construct a canonical frame from an orthonormal frame as follows.

Proposition A.2 (see [Le2, Proposition 3.2]) *Given a frame $\bar{E}^t = (\bar{e}_i^t)_{i=1}^n$ orthonormal in the sense of Lemma A.1, put $\Omega_{ij}^t := \omega(\dot{\bar{e}}_i^t, \dot{\bar{e}}_j^t)$ and let O^t be the solution to*

$$\dot{O}^t = \frac{1}{2} O^t \Omega^t, \quad O^0 = I_n.$$

Then $E^t = (e_i^t)_{i=1}^n$ with $e_i^t := \sum_{j=1}^n O_{ij}^t \bar{e}_j^t$ gives a canonical frame.

Proof. We will sometimes omit the evaluations at time t for brevity. First of all, O^t is orthogonal for all t since

$$(OO^T)' = \frac{1}{2} O \Omega O^T + \frac{1}{2} O \Omega^T O^T = \frac{1}{2} O (\Omega + \Omega^T) O^T = 0,$$

where O^T denotes the transpose of O . To see that E^t is orthonormal, observe

$$\omega(\dot{E}, E) = \frac{1}{2} \omega(O \Omega \bar{E}, O \bar{E}) + \omega(O \dot{\bar{E}}, O \bar{E}) = \frac{1}{2} O \Omega \omega(\bar{E}, \bar{E}) O^T + O \omega(\dot{\bar{E}}, \bar{E}) O^T.$$

Then it follows from (the proof of) Lemma 3.5 and $\omega(\dot{\bar{E}}, \bar{E}) \equiv I_n$ that $\omega(\dot{E}^t, E^t) \equiv I_n$.

It remains to show $\dot{e}_i^t \in \mathcal{J}_\alpha^t$. We have

$$\ddot{E} = \left(\frac{1}{2} O \Omega \bar{E} + O \dot{\bar{E}} \right)' = \frac{1}{2} (O \Omega)' \bar{E} + O \Omega \dot{\bar{E}} + O \ddot{\bar{E}},$$

and hence

$$\omega(\ddot{E}, \bar{E}) = \omega(O \Omega \dot{\bar{E}}, \bar{E}) + \omega(O \ddot{\bar{E}}, \bar{E}) = O \Omega + O \omega(\ddot{\bar{E}}, \bar{E}) = O \left[\omega(\dot{\bar{E}}, \bar{E}) \right]' \equiv 0.$$

This implies $\dot{e}_i^t \in \mathcal{J}_\alpha^t$ because the intersection of the kernels of $\omega(\cdot, \bar{e}_j^t)$, $j = 1, \dots, n$, coincides with \mathcal{J}_α^t . □

A.2 Coordinate representation

We compute the curvature operator in coordinates by using the canonical frame in the previous subsection. First of all, Proposition A.2 yields the following (see [Le2, Proposition 3.3]):

$$\begin{aligned}\mathbf{R}_\alpha^0(E^0) &= -(O^t \bar{E}^t)''|_{t=0} = -\left(\frac{1}{2}O^t \Omega^t \bar{E}^t + O^t \dot{\bar{E}}^t\right)' \Big|_{t=0} \\ &= -\frac{1}{4}(\Omega^0)^2 \bar{E}^0 - \frac{1}{2}\dot{\Omega}^0 \bar{E}^0 - \Omega^0 \dot{\bar{E}}^0 - \ddot{\bar{E}}^0.\end{aligned}\quad (\text{A.1})$$

We choose a coordinate satisfying $\mathbf{H}_{\alpha_i \alpha_j}(\alpha) = \delta_{ij}$ for all i, j , so that we can take $\xi_i(0) = \partial_{\alpha_i}$. Then it holds that, by the calculation in Lemma A.1,

$$\ddot{e}_i^0 = -\partial_{x^i} + \sum_{k=1}^n \{\mathbf{H}_{x^k \alpha_i}(\alpha) + \dot{\xi}_i^k(0)\} \partial_{\alpha_k}.\quad (\text{A.2})$$

This implies

$$\Omega_{ij}^0 = \omega(\ddot{e}_i^0, \ddot{e}_j^0) = \mathbf{H}_{x^i \alpha_j}(\alpha) - \mathbf{H}_{x^j \alpha_i}(\alpha) + \dot{\xi}_j^i(0) - \dot{\xi}_i^j(0).\quad (\text{A.3})$$

Now, for further simplification, let us assume $\mathbf{H}_{\alpha_i \alpha_j}(\alpha(t)) = \delta_{ij}$ and put $\xi_i(t) = \partial_{\alpha_i}$ for all t . It is always possible to choose such a coordinate by noticing that

$$\left(\sum_{i=1}^n a_i \partial_{x^i}, \sum_{j=1}^n b_j \partial_{x^j}\right) \mapsto \sum_{i,j=1}^n a_i b_j \mathbf{L}_{v^i v^j}(\dot{\eta}(t)), \quad \eta(t) := \pi_M(\alpha(t)),$$

gives an inner product of $T_{\eta(t)}M$ for each t and by taking $(x^i)_{i=1}^n$ such that $(\partial_{x^i})_{i=1}^n$ is orthonormal with respect to this inner product at every $\eta(t)$. Then we have, by omitting evaluations at α and $t = 0$ for brevity,

$$\begin{aligned}\ddot{e}_i^0 &= -\sum_{k=1}^n \{[\mathbf{H}_{\alpha_k \alpha_i} \circ \alpha]' \partial_{x^k} - [\mathbf{H}_{x^k \alpha_i} \circ \alpha]' \partial_{\alpha_k}\} \\ &\quad + \sum_{k,l=1}^n \{\mathbf{H}_{\alpha_k \alpha_i}(\mathbf{H}_{\alpha_l x^k} \partial_{x^l} - \mathbf{H}_{x^l x^k} \partial_{\alpha_l}) - \mathbf{H}_{x^k \alpha_i}(\mathbf{H}_{\alpha_l \alpha_k} \partial_{x^l} - \mathbf{H}_{x^l \alpha_k} \partial_{\alpha_l})\} \\ &= \sum_{k=1}^n (\mathbf{H}_{\alpha_k x^i} - \mathbf{H}_{x^k \alpha_i}) \partial_{x^k} + \sum_{k=1}^n \left\{ [\mathbf{H}_{x^k \alpha_i} \circ \alpha]' - \mathbf{H}_{x^k x^i} + \sum_{l=1}^n \mathbf{H}_{x^l \alpha_i} \mathbf{H}_{x^k \alpha_l} \right\} \partial_{\alpha_k}.\end{aligned}$$

Combining this with (A.1), (A.2) and (A.3), we see that the horizontal part of $\mathbf{R}_\alpha^0(e_i^0)$ indeed vanishes (so that $\mathbf{R}_\alpha^0(e_i^0) \in \mathcal{V}_\alpha$). We finally calculate

$$\begin{aligned}\dot{\Omega}_{ij}^0 &= \omega(\dot{e}_i^0, \dot{e}_j^0) + \omega(\ddot{e}_i^0, \dot{e}_j^0) \\ &= [\mathbf{H}_{x^i \alpha_j} \circ \alpha]' - \mathbf{H}_{x^i x^j} + \sum_{k=1}^n \mathbf{H}_{x^k \alpha_j} \mathbf{H}_{x^i \alpha_k} + \sum_{k=1}^n \mathbf{H}_{x^k \alpha_i} (\mathbf{H}_{\alpha_k x^j} - \mathbf{H}_{x^k \alpha_j}) \\ &\quad - [\mathbf{H}_{x^j \alpha_i} \circ \alpha]' + \mathbf{H}_{x^j x^i} - \sum_{k=1}^n \mathbf{H}_{x^k \alpha_i} \mathbf{H}_{x^j \alpha_k} - \sum_{k=1}^n \mathbf{H}_{x^k \alpha_j} (\mathbf{H}_{\alpha_k x^i} - \mathbf{H}_{x^k \alpha_i}) \\ &= [(\mathbf{H}_{x^i \alpha_j} - \mathbf{H}_{x^j \alpha_i}) \circ \alpha]'.\end{aligned}$$

Consequently, we obtain that the coefficient of $\partial_{\alpha_j} = e_j^0$ in $\mathbf{R}_\alpha^0(e_i^0)$ is

$$\begin{aligned}
& -\frac{1}{4} \sum_{k=1}^n (\mathbf{H}_{x^i \alpha_k} - \mathbf{H}_{x^k \alpha_i})(\mathbf{H}_{x^k \alpha_j} - \mathbf{H}_{x^j \alpha_k}) - \frac{1}{2} [(\mathbf{H}_{x^i \alpha_j} - \mathbf{H}_{x^j \alpha_i}) \circ \boldsymbol{\alpha}]' \\
& - \sum_{k=1}^n (\mathbf{H}_{x^i \alpha_k} - \mathbf{H}_{x^k \alpha_i}) \mathbf{H}_{x^j \alpha_k} - [\mathbf{H}_{x^j \alpha_i} \circ \boldsymbol{\alpha}]' + \mathbf{H}_{x^j x^i} - \sum_{k=1}^n \mathbf{H}_{x^k \alpha_i} \mathbf{H}_{x^j \alpha_k} \\
& = \frac{1}{4} \sum_{k=1}^n (\mathbf{H}_{x^i \alpha_k} - \mathbf{H}_{x^k \alpha_i})(\mathbf{H}_{x^j \alpha_k} - \mathbf{H}_{x^k \alpha_j}) - \frac{1}{2} [(\mathbf{H}_{x^i \alpha_j} + \mathbf{H}_{x^j \alpha_i}) \circ \boldsymbol{\alpha}]' \\
& - \sum_{k=1}^n \mathbf{H}_{x^i \alpha_k} \mathbf{H}_{x^j \alpha_k} + \mathbf{H}_{x^i x^j}. \tag{A.4}
\end{aligned}$$

Note that this is indeed symmetric in i and j .

References

- [Agr1] A. A. Agrachev, Curvature and hyperbolicity of Hamiltonian systems (Russian), Tr. Mat. Inst. Steklova **256** (2007), Din. Sist. i Optim., 31–53; translation in Proc. Steklov Inst. Math. **256** (2007), 26–46.
- [Agr2] A. A. Agrachev, Geometry of optimal control problems and Hamiltonian systems, Non-linear and optimal control theory, 1–59, Lecture Notes in Math., **1932**, Springer, Berlin, 2008.
- [AG] A. A. Agrachev and R. V. Gamkrelidze, Feedback-invariant optimal control theory and differential geometry. I. Regular extremals, J. Dynam. Control Systems **3** (1997), 343–389.
- [AL1] A. A. Agrachev and P. W. Y. Lee, Generalized Ricci curvature bounds for three dimensional contact subriemannian manifolds, Preprint (2009). Available at [arXiv:0903.2550](https://arxiv.org/abs/0903.2550)
- [AL2] A. A. Agrachev and P. W. Y. Lee, Bishop and Laplacian comparison theorems on three dimensional contact subriemannian manifolds with symmetry, to appear in J. Geom. Anal. Available at [arXiv:1105.2206](https://arxiv.org/abs/1105.2206)
- [Agu] M. Agueh, Finsler structure in the p -Wasserstein space and gradient flows, C. R. Math. Acad. Sci. Paris **350** (2012), 35–40.
- [AD] J. C. Álvarez Paiva and C. E. Durán, Geometric invariants of fanning curves, Adv. in Appl. Math. **42** (2009), 290–312.
- [AGS1] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures, Birkhäuser Verlag, Basel, 2005.
- [AGS2] L. Ambrosio, N. Gigli and G. Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces, Rev. Mat. Iberoam. **29** (2013), 969–996.
- [AGS3] L. Ambrosio, N. Gigli and G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, Invent. Math. **195** (2014), 289–391.

- [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.
- [BB] P. Bernard and B. Buffoni, Optimal mass transportation and Mather theory, J. Eur. Math. Soc. (JEMS) **9** (2007), 85–121.
- [Bre] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. **44** (1991), 375–417.
- [Bré] H. Brézis, Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert (French), North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.
- [CMS] D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschläger, A Riemannian interpolation inequality à 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.
- [Er] M. Erbar, The heat equation on manifolds as a gradient flow in the Wasserstein space, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), 1–23.
- [FF] A. Fathi and A. Figalli, Optimal transportation on non-compact manifolds, Israel J. Math. **175** (2010), 1–59.
- [Fo] P. Foulon, Géométrie des équations différentielles du second ordre (French), Ann. Inst. H. Poincaré Phys. Théor. **45** (1986), 1–28.
- [FGM] N. Fusco, M. Gori and F. Maggi, A remark on Serrin’s theorem, NoDEA Nonlinear Differential Equations Appl. **13** (2006), 425–433.
- [Gi] N. Gigli, On the differential structure of metric measure spaces and applications, Preprint (2012). Available at [arXiv:1205.6622](https://arxiv.org/abs/1205.6622)
- [GKO] N. Gigli, K. Kuwada and S. Ohta, Heat flow on Alexandrov spaces, Comm. Pure Appl. Math. **66** (2013), 307–331.
- [GO] N. Gigli and S. Ohta, First variation formula in Wasserstein spaces over compact Alexandrov spaces, Canad. Math. Bull. **55** (2012), 723–735.
- [Gr] J. Grifone, Structure presque-tangente et connexions. I (French), Ann. Inst. Fourier (Grenoble) **22** (1972), 287–334.
- [JKO] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker–Planck equation, SIAM J. Math. Anal. **29** (1998), 1–17.
- [Ju] N. Juillet, Geometric inequalities and generalized Ricci bounds in the Heisenberg group, Int. Math. Res. Not. IMRN (2009), 2347–2373.
- [Ke1] M. Kell, On interpolation and curvature via Wasserstein geodesics, Preprint (2013). Available at [arXiv:1311.5407](https://arxiv.org/abs/1311.5407)

- [Ke2] M. Kell, q -heat flow and the gradient flow of the Renyi entropy in the p -Wasserstein space, Preprint (2014). Available at [arXiv:1401.0840](https://arxiv.org/abs/1401.0840)
- [Le1] P. W. Y. Lee, A remark on the potentials of optimal transport maps, *Acta Appl. Math.* **115** (2011), 123–138.
- [Le2] P. W. Y. Lee, Displacement interpolations from a Hamiltonian point of view, *J. Funct. Anal.* **265** (2013), 3163–3203.
- [LLZ] P. W. Y. Lee, C. Li and I. Zelenko, Measure contraction properties of contact sub-Riemannian manifolds with symmetry, Preprint (2013). Available at [arXiv:1304.2658](https://arxiv.org/abs/1304.2658)
- [LV1] J. Lott and C. Villani, Weak curvature conditions and functional inequalities, *J. Funct. Anal.* **245** (2007), 311–333.
- [LV2] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport, *Ann. of Math. (2)* **169** (2009), 903–991.
- [Maa] J. Maas, Gradient flows of the entropy for finite Markov chains, *J. Funct. Anal.* **261** (2011), 2250–2292.
- [May] U. F. Mayer, Gradient flows on nonpositively curved metric spaces and harmonic maps, *Comm. Anal. Geom.* **6** (1998), 199–253.
- [Mc] R. J. McCann, Polar factorization of maps on Riemannian manifolds, *Geom. Funct. Anal.* **11** (2001), 589–608.
- [Oh1] S. Ohta, On the measure contraction property of metric measure spaces, *Comment. Math. Helv.* **82** (2007), 805–828.
- [Oh2] S. Ohta, Gradient flows on Wasserstein spaces over compact Alexandrov spaces, *Amer. J. Math.* **131** (2009), 475–516.
- [Oh3] S. Ohta, Uniform convexity and smoothness, and their applications in Finsler geometry, *Math. Ann.* **343** (2009), 669–699.
- [Oh4] S. Ohta, Finsler interpolation inequalities, *Calc. Var. Partial Differential Equations* **36** (2009), 211–249.
- [Oh5] S. Ohta, Vanishing S-curvature of Randers spaces, *Differential Geom. Appl.* **29** (2011), 174–178.
- [Oh6] S. Ohta, Splitting theorems for Finsler manifolds of nonnegative Ricci curvature, to appear in *J. Reine Angew. Math.* Available at [arXiv:1203.0079](https://arxiv.org/abs/1203.0079)
- [OS1] S. Ohta and K.-T. Sturm, Heat flow on Finsler manifolds, *Comm. Pure Appl. Math.* **62** (2009), 1386–1433.
- [OS2] S. Ohta and K.-T. Sturm, Non-contraction of heat flow on Minkowski spaces, *Arch. Ration. Mech. Anal.* **204** (2012), 917–944.
- [OS3] S. Ohta and K.-T. Sturm, Bochner–Weitzenböck formula and Li–Yau estimates on Finsler manifolds, *Adv. Math.* **252** (2014), 429–448.

- [OT1] S. Ohta and A. Takatsu, Displacement convexity of generalized relative entropies, *Adv. Math.* **228** (2011), 1742–1787.
- [OT2] S. Ohta and A. Takatsu, Displacement convexity of generalized relative entropies. II, *Comm. Anal. Geom.* **21** (2013), 687–785.
- [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), 923–940.
- [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.
- [Se] J. Serrin, On the definition and properties of certain variational integrals, *Trans. Amer. Math. Soc.* **101** (1961), 139–167.
- [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, On the geometry of metric measure spaces. I, *Acta Math.* **196** (2006), 65–131.
- [St2] K.-T. Sturm, On the geometry of metric measure spaces. II, *Acta Math.* **196** (2006), 133–177.
- [Vi] C. Villani, *Optimal transport, old and new*, Springer-Verlag, Berlin, 2009.
- [WX] G. Wang and C. Xia, A sharp lower bound for the first eigenvalue on Finsler manifolds, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **30** (2013), 983–996.
- [WW] G. Wei and W. Wylie, Comparison geometry for the Bakry-Emery Ricci tensor, *J. Differential Geom.* **83** (2009), 377–405.
- [Xi] C. Xia, Local gradient estimate for harmonic functions on Finsler manifolds, to appear in *Calc. Var. Partial Differential Equations*. Available at [arXiv:1308.3609](https://arxiv.org/abs/1308.3609)