A semigroup approach to Finsler geometry: Bakry–Ledoux’s isoperimetric inequality

Shin-ichi Ohta

June 23, 2020

Abstract

We develop the celebrated semigroup approach à la Bakry et al on Finsler manifolds, where natural Laplacian and heat semigroup are nonlinear, based on the Bochner–Weitzenböck formula established by Sturm and the author. We show the $L^1$-gradient estimate on Finsler manifolds (under some additional assumptions in the noncompact case), which is equivalent to a lower weighted Ricci curvature bound and the improved Bochner inequality. As a geometric application, we prove Bakry–Ledoux’s Gaussian isoperimetric inequality, again under some additional assumptions in the noncompact case. This extends Cavalletti–Mondino’s inequality on reversible Finsler manifolds to non-reversible metrics, and improves the author’s previous estimate, both based on the localization (also called needle decomposition) method.

Mathematics Subject Classification (2010): 53C60, 58J35, 49Q20

Contents

1 Introduction 2

2 Geometry and analysis on Finsler manifolds 4
  2.1 Finsler manifolds ........................................ 5
  2.2 Uniform convexity and smoothness ...................... 7
  2.3 Weighted Ricci curvature .................................. 9
  2.4 Nonlinear Laplacian and heat flow ...................... 10
  2.5 Bochner–Weitzenböck formula .......................... 12

3 Linearized semigroups and gradient estimates 13
  3.1 Linearized heat semigroups and their adjoints .......... 13
  3.2 Improved Bochner inequality ............................ 16
  3.3 $L^1$-gradient estimate ................................. 17
1 Introduction

The aim of this article is to put forward the semigroup approach in geometric analysis on Finsler manifolds, based on the Bochner–Weitzenböck formula established in [OS3]. There are already a number of applications of the Bochner–Weitzenböck formula (including [WX, Xi, YH, Oh7]), and the machinery in this article would contribute to a further development. In addition, our treatment of a nonlinear generator and the associated nonlinear semigroup (Laplacian and heat semigroup) could be of independent interest from the analytic viewpoint.

The celebrated theory developed by Bakry, Emery, Ledoux et al (called the $\Gamma$-calculus) studies symmetric generators and the associated linear, symmetric diffusion semigroups under a kind of Bochner inequality (called the (analytic) curvature-dimension condition). Attributed to Bakry–Emery’s original work [BE], this condition will be denoted by $\text{BE}(K, N)$ in this introduction, where $K \in \mathbb{R}$ and $N \in (1, \infty]$ are parameters corresponding to ‘curvature’ and ‘dimension’, respectively. This technique is extremely powerful in studying various inequalities (log-Sobolev and Poincaré inequalities, gradient estimates, etc.) in a unified way, we refer to [BE] and the recent book [BGL] for more on this theory.

On a Riemannian manifold equipped with the Laplacian $\Delta$, $\text{BE}(K, N)$ means the following Bochner-type inequality:

$$\Delta \left[ \frac{||\nabla u||^2}{2} \right] - \langle \nabla (\Delta u), \nabla u \rangle \geq K ||\nabla u||^2 + \frac{(\Delta u)^2}{N}.$$ 

Thereby a Riemannian manifold with Ricci curvature not less than $K$ and dimension not greater than $N$ (more generally, a weighted Riemannian manifold of weighted Ricci curvature $\text{Ric}_N \geq K$) is a fundamental example satisfying $\text{BE}(K, N)$.

Later, inspired by [CMS, OV], Sturm [vRS, St1, St2] and Lott–Villani [LV] introduced the (geometric) curvature-dimension condition $\text{CD}(K, N)$ for metric measure spaces in terms of optimal transport theory. The condition $\text{CD}(K, N)$ characterizes $\text{Ric} \geq K$ and $\text{dim} \leq N$ (or $\text{Ric}_N \geq K$) for (weighted) Riemannian manifolds, and its formulation requires a lower regularity of spaces than $\text{BE}(K, N)$. We refer to Villani’s book [Vi] for more on this rapidly developing theory. It was shown in [Oh2] that $\text{CD}(K, N)$ also holds and characterizes $\text{Ric}_N \geq K$ for Finsler manifolds, where the natural Laplacian and the associated heat semigroup are nonlinear. For this reason, Ambrosio, Gigli and Savaré [AGS1] introduced a reinforced version $\text{RCD}(K, \infty)$ called the Riemannian curvature-dimension condition as the combination of $\text{CD}(K, \infty)$ and the linearity of heat.
semigroup, followed by the finite-dimensional analogue $\text{RCD}^*(K, N)$ investigated by Erbar, Kuwada and Sturm [EKS] (see also [Gi1, Gi2]). It then turned out that $\text{RCD}^*(K, N)$ is equivalent to $\text{BE}(K, N)$ ([AGS2, EKS]), this equivalence justifies the term ‘curvature-dimension condition’ which actually came from the similarity to Bakry’s theory.

In this article, we develop the theory of Bakry et al on Finsler manifolds. We consider a Finsler manifold $M$ equipped with a Finsler metric $F : TM \to [0, \infty)$ and a positive $C^\infty$-measure $m$ on $M$. We will not assume that $F$ is reversible, thereby $F(-v) \neq F(v)$ is allowed. The key ingredient, the Bochner inequality under $\text{Ric}_N \geq K$, was established in [OS3] as follows:

$$
\Delta u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) \geq K F^2(\nabla u) + \frac{(\Delta u)^2}{N}. \tag{1.1}
$$

This Bochner inequality has the same form as the Riemannian case by means of the mixture of the nonlinear Laplacian $\Delta$ and its linearization $\Delta F^u$. Despite of this mixture, we could derive Bakry–Emery’s $L^2$-gradient estimate as well as Li–Yau’s estimates on compact manifolds (see [OS3, §4]). We proceed further in this direction and show the improved Bochner inequality under $\text{Ric}_\infty \geq K$ (Proposition 3.5):

$$
\Delta u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) - K F^2(\nabla u) \geq d[F(\nabla u)](\nabla F^u[F(\nabla u)]). \tag{1.2}
$$

The first application of (1.2) is the $L^1$-gradient estimate (Theorem 3.7), where we include also the noncompact case but with some additional (likely redundant) assumptions, see the theorem below where we assume the same conditions. We also see that the Bochner inequalities (1.1) (with $N = \infty$), (1.2) and the $L^2$- and $L^1$-gradient estimates are all equivalent to $\text{Ric}_\infty \geq K$ (Theorem 3.9).

The second, geometric application of (1.2) is a generalization of Bakry–Ledoux’s Gaussian isoperimetric inequality (Theorem 4.1):

**Theorem (Bakry–Ledoux’s isoperimetric inequality)** Let $(M, F, m)$ be complete and satisfy $\text{Ric}_\infty \geq K > 0$, $m(M) = 1$, $C_F < \infty$ and $S_F < \infty$. We also assume that

$$
d[F(\nabla u_t)](\nabla F^u[F(\nabla u_t)]) \in L^1(M)
$$

holds for any global solution $(u_t)_{t \geq 0}$ to the heat equation with $u_0 \in C^\infty_c(M)$ and any $t > 0$. Then we have

$$
\mathcal{I}_{(M, F, m)}(\theta) \geq \mathcal{I}_K(\theta) \tag{1.3}
$$

for all $\theta \in [0, 1]$, where

$$
\mathcal{I}_K(\theta) := \sqrt{\frac{K}{2\pi}} e^{-Kc^2(\theta)/2} \quad \text{with} \quad \theta = \int_{-\infty}^{c(\theta)} \sqrt{\frac{K}{2\pi}} e^{-Kc^2/2} \, da.
$$

Here $\mathcal{I}_{(M, F, m)} : [0, 1] \to [0, \infty)$ is the isoperimetric profile defined as the least boundary area of sets $A \subset M$ with $m(A) = \theta$ (see the beginning of Section 4), and $C_F$ (resp. $S_F$) is the $(2)$-uniform convexity (resp. smoothness) constant which bounds the reversibility,

$$
\Lambda_F := \sup_{v \in TM \setminus \{0\}} \frac{F(v)}{F(-v)} \in [1, \infty], \tag{1.4}
$$
as $\Lambda_F \leq \min\{\sqrt{C_F}, \sqrt{S_F}\}$ (see Lemma 2.4). (In particular, the forward completeness is equivalent to the backward completeness, and we denoted it by the plain completeness in the theorem.) All the conditions $C_F < \infty$, $S_F < \infty$, and $d[F(\nabla u_t)](\nabla \nabla \nabla [F(\nabla u_t)]) \in L^1(M)$ hold true in the compact case. In the noncompact case, however, there are technical difficulties and it is unclear how to remove them in this semigroup approach (see §3.4 for a discussion). We remark that, in [Oh8] based on the needle decomposition, we did not need those conditions.

The inequality (1.3) has the same form as the Riemannian case in [BL], and it is sharp and the model space is the real line $\mathbb{R}$ equipped with the normal (Gaussian) distribution $d\mu = \sqrt{K/2\pi} e^{-Kx^2/2} dx$. See [BL] for the original work of Bakry and Ledoux on linear diffusion semigroups (influenced by Bobkov’s works [Bob1, Bob2]), and [Bor, SC] for the classical Euclidean or Hilbert cases. We also refer to [AM] for the Gaussian isoperimetric inequality on RCD($K, \infty$)-spaces by a refinement of the $\Gamma$-calculus.

The above theorem extends Cavalletti–Mondino’s isoperimetric inequality in [CM] to non-reversible Finsler manifolds. Precisely, they considered essentially non-branching metric measure spaces $(X, d, m)$ satisfying CD($K, N$) for $K \in \mathbb{R}$ and $N \in (1, \infty)$, and showed the sharp Lévy–Gromov type isoperimetric inequality of the form

$$I_{(X,d,m)}(\theta) \geq I_{K,N,D}(\theta)$$

with diam $X \leq D$ ($\leq \infty$). The case of $N = \infty$ is not included in [CM] for technical reasons on the structure of CD($K, \infty$)-spaces, but the same argument gives (1.3) (corresponding to $N = D = \infty$) for reversible Finsler manifolds. The proof in [CM] is based on the needle decomposition (also called localization) inspired by Klartag’s work [Kl] on Riemannian manifolds, extending the successful technique in convex geometry. Along the lines of [CM], in [Oh8] we have generalized the needle decomposition to non-reversible Finsler manifolds, however, then we obtain only a weaker isoperimetric inequality,

$$I_{(M,F,m)}(\theta) \geq \Lambda_F^{-1} \cdot I_{K,N,D}(\theta),$$

with $\Lambda_F$ in (1.4). The inequality (1.3) improves (1.5) in the case where $N = D = \infty$ and $K > 0$, and supports a conjecture that the sharp isoperimetric inequality in the non-reversible case is the same as the reversible case, namely $\Lambda_F^{-1}$ in (1.5) would be removed.

The organization of this article is as follows: In Section 2 we review the basics of Finsler geometry, including the weighted Ricci curvature and the Bochner–Weitzenböck formula. Section 3 is devoted to a detailed study of the nonlinear heat semigroup and its linearizations, we improve the Bochner inequality under $\text{Ric}_{\infty} \geq K$ and show the $L^1$-gradient estimate. We prove the isoperimetric inequality in Section 4.

Acknowledgements. I am grateful to Kazumasa Kuwada for his suggestion to consider this problem and for many valuable discussions. I also thank Karl-Theodor Sturm and Kohei Suzuki for stimulating discussions.

## 2 Geometry and analysis on Finsler manifolds

We review the basics of Finsler geometry (we refer to [BCS, Sh] for further reading), and introduce the weighted Ricci curvature and the nonlinear Laplacian studied in [Oh2, OS1]
Throughout the article, let \( M \) be a connected \( C^\infty \)-manifold without boundary of dimension \( n \geq 2 \). We also fix an arbitrary positive \( C^\infty \)-measure \( m \) on \( M \).

### 2.1 Finsler manifolds

Given local coordinates \( (x^i)_{i=1}^n \) on an open set \( U \subset M \), we will always use the fiber-wise linear coordinates \( (x^i, v^j)_{i,j=1}^n \) of \( TU \) such that

\[
v = \sum_{j=1}^n v^j \frac{\partial}{\partial x^j} \bigg|_x \in T_x M, \quad x \in U.
\]

**Definition 2.1 (Finsler structures)** We say that a nonnegative function \( F : TM \rightarrow [0, \infty) \) is a \( C^1 \)-Finsler structure of \( M \) if the following three conditions hold:

1. **(Regularity)** \( F \) is \( C^1 \) on \( TM \setminus 0 \), where 0 stands for the zero section;

2. **(Positive 1-homogeneity)** It holds \( F(cv) = cF(v) \) for all \( v \in TM \) and \( c \geq 0 \);

3. **(Strong convexity)** The \( n \times n \) matrix

\[
\left(g_{ij}(v)\right)^n_{i,j=1} := \left(\frac{1}{2} \frac{\partial^2 (F^2)}{\partial v^i \partial v^j}(v)\right)^n_{i,j=1}
\]

is positive-definite for all \( v \in TM \setminus 0 \).

We call such a pair \((M, F)\) a \( C^\infty \)-Finsler manifold.

In other words, \( F \) provides a Minkowski norm on each tangent space which varies smoothly in horizontal directions. If \( F(-v) = F(v) \) holds for all \( v \in TM \), then we say that \( F \) is reversible or absolutely homogeneous. The strong convexity means that the unit sphere \( T_x M \setminus F^{-1}(1) \) (called the indicatrix) is ‘positively curved’ and implies the strict convexity: \( F(v + w) \leq F(v) + F(w) \) for all \( v, w \in T_x M \) and equality holds only when \( v = aw \) or \( w = av \) for some \( a \geq 0 \).

In the coordinates \( (x^i, \alpha_j)_{i,j=1}^n \) of \( T^*U \) given by \( \alpha = \sum_{j=1}^n \alpha_j \, dx^j \), we will also consider

\[
g_{ij}(\alpha) := \frac{1}{2} \frac{\partial^2 [F^*]^{2}}{\partial \alpha_i \partial \alpha_j}(\alpha), \quad i, j = 1, 2, \ldots, n,
\]

for \( \alpha \in T^* U \setminus 0 \). Here \( F^* : T^*_x M \rightarrow [0, \infty) \) is the dual Minkowski norm to \( F \), namely

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

for \( \alpha \in T^*_x M \). It is clear by definition that \( \alpha(v) \leq F^*(\alpha)F(v) \), and hence

\[
\alpha(v) \geq -F^*(\alpha)F(-v), \quad \alpha(v) \geq -F^*(-\alpha)F(v).
\]

We remark that, however, \( \alpha(v) \geq -F^*(\alpha)F(v) \) does not hold in general.
Let us denote by $\mathcal{L}^*: T^*M \rightarrow TM$ the Legendre transform. Precisely, $\mathcal{L}^*$ is sending $\alpha \in T^*_x M$ to the unique element $v \in T_x M$ such that $F(v) = F^*(\alpha)$ and $\alpha(v) = F^*(\alpha)^2$. In coordinates we can write down

$$\mathcal{L}^*(\alpha) = \sum_{i,j=1}^n g^{ij}_x(\alpha) \partial / \partial x^j \bigg|_x = \sum_{j=1}^n \frac{1}{2} \frac{\partial [(F^*)^2]}{\partial \alpha_j}(\alpha) \frac{\partial}{\partial x^j} \bigg|_x$$

for $\alpha \in T^*_x M \setminus 0$ (the latter expression makes sense also at 0). Note that $g^{ij}_x(\alpha) = g^{ij}(\mathcal{L}^*(\alpha))$ for $\alpha \in T^*_x M \setminus 0$, where $(g^{ij}(v))$ denotes the inverse matrix of $(g_{ij}(v))$. The map $\mathcal{L}^*: T^*_x M$ is being a linear operator only when $F|_{T_x M}$ comes from an inner product. We also define $\mathcal{L} := (\mathcal{L}^*)^{-1}: TM \rightarrow T^*_x M$.

For $x, y \in M$, we define the (asymmetric) distance from $x$ to $y$ by

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

where $\eta: [0, 1] \rightarrow M$ runs over all $C^1$-curves such that $\eta(0) = x$ and $\eta(1) = y$. Note that $d(y, x) \neq d(x, y)$ can happen since $F$ is only positively homogeneous. A $C^\infty$-curve $\eta$ on $M$ is called a geodesic if it is locally minimizing and has a constant speed with respect to $d$, similarly to Riemannian or metric geometry. See (2.7) below for the precise geodesic equation. For $v \in T_x M$, if there is a geodesic $\eta: [0, 1] \rightarrow M$ with $\dot{\eta}(0) = v$, then we define the exponential map by $\exp_x(v) := \eta(1)$. We say that $(M, F)$ is forward complete if the exponential map is defined on whole $TM$. Then the Hopf-Rinow theorem ensures that any pair of points is connected by a minimal geodesic (see [BCS, Theorem 6.6.1]).

Given each $v \in T_x M \setminus 0$, the positive-definite matrix $(g_{ij}(v))_{i,j=1}^n$ in (2.1) induces the Riemannian structure $g_v$ of $T_x M$ by

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

Notice that this definition is coordinate-free and $g_v(v, v) = F^2(v)$ holds. One can regard $g_v$ as the best Riemannian approximation of $F|_{T_x M}$ in the direction $v$. The Cartan tensor

$$A_{ijk}(v) := \frac{F(v)}{2} \frac{\partial g_{ij}}{\partial v^k}(v), \quad v \in TM \setminus 0,$$

measures the variation of $g_v$ in vertical directions, and vanishes everywhere on $TM \setminus 0$ if and only if $F$ comes from a Riemannian metric.

The following useful fact on homogeneous functions (see [BCS, Theorem 1.2.1]) plays a fundamental role in our calculus.

**Theorem 2.2 (Euler’s theorem)** Suppose that a differentiable function $H: \mathbb{R}^n \setminus 0 \rightarrow \mathbb{R}$ satisfies $H(cv) = c^r H(v)$ for some $r \in \mathbb{R}$ and all $c > 0$ and $v \in \mathbb{R}^n \setminus 0$ (that is, positively $r$-homogeneous). Then we have, for all $v \in \mathbb{R}^n \setminus 0$,

$$\sum_{i=1}^n \frac{\partial H}{\partial v^i}(v) v^i = r H(v).$$
Observe that $g_{ij}$ is positively 0-homogeneous on each $T_xM$, and hence

$$\sum_{i=1}^{n} A_{ijk}(v)v^i = \sum_{j=1}^{n} A_{ijk}(v)v^j = \sum_{k=1}^{n} A_{ijk}(v)v^k = 0 \quad (2.3)$$

for all $v \in TM \setminus 0$ and $i, j, k = 1, 2, \ldots, n$. Define the formal Christoffel symbol

$$\gamma_{jk}^i(v) := \frac{1}{2} \sum_{l,m=1}^{n} g^{il}(v) \left\{ \frac{\partial g_{lk}}{\partial x^j}(v) + \frac{\partial g_{jl}}{\partial x^k}(v) - \frac{\partial g_{jk}}{\partial x^l}(v) \right\} \quad (2.4)$$

for $v \in TM \setminus 0$, and the geodesic spray coefficients and the nonlinear connection

$$G^i(v) := \sum_{j,k=1}^{n} \gamma_{jk}^i(v)v^jv^k, \quad N_j^i(v) := \frac{1}{2} \partial G^i \bigg|_{v^j}$$

for $v \in TM \setminus 0$ ($G^i(0) = N_j^i(0) := 0$ by convention). Note that $G^i$ is positively 2-homogeneous, hence Theorem 2.2 implies $\sum_{j=1}^{n} N_j^i(v)v^j = G^i(v)$.

By using $N_j^i$, the coefficients of the Chern connection are given by

$$\Gamma_{jk}^i(v) := \gamma_{jk}^i(v) - \sum_{l,m=1}^{n} \frac{g^{il}}{F} \left( A_{lkm}N_j^m + A_{jlm}N_k^m - A_{jkm}N_l^m \right) (v) \quad (2.5)$$

on $TM \setminus 0$. The corresponding covariant derivative of a vector field $X$ by $v \in T_xM$ with reference vector $w \in T_xM \setminus 0$ is defined as

$$D^w_vX(x) := \sum_{i,j=1}^{n} \left\{ v^j \frac{\partial X^i}{\partial x^j}(x) + \sum_{k=1}^{n} \Gamma_{jk}^i(w)v^jX^k(x) \right\} \frac{\partial}{\partial x^i} \bigg|_x \in T_xM. \quad (2.6)$$

Then the geodesic equation is written as, with the help of (2.3),

$$D^w_v\dot{q}(t) = \sum_{i=1}^{n} \left\{ \dot{q}^i(t) + G^i(\dot{q}(t)) \right\} \frac{\partial}{\partial x^i} \bigg|_{\eta(t)} = 0. \quad (2.7)$$

### 2.2 Uniform convexity and smoothness

We will need the following quantity associated with $(M, F)$:

$$S_F := \sup_{x \in M} \sup_{v, w \in T_xM \setminus 0} \frac{g_v(w, w)}{F^2(w)} \in [1, \infty].$$

Since $g_v(w, w) \leq S_f F^2(w)$ and $g_v$ is the Hessian of $F^2/2$ at $v$, the constant $S_F$ measures the (fiber-wise) concavity of $F^2$ and is called the (2-)uniform smoothness constant (see [Oh1]). We remark that $S_F = 1$ holds if and only if $(M, F)$ is Riemannian. The following lemma is a standard fact, we give a proof for thoroughness.
Lemma 2.3 For any $x \in M$, $v \in T_x M \setminus 0$ and $\alpha := \mathcal{L}(v)$, we have
\[
\sup_{w \in T_x M \setminus 0} \frac{g_v(w, w)}{F^2(w)} = \sup_{\beta \in T_x^* M \setminus 0} \frac{F^*(\beta)^2}{g^*_\alpha(\beta, \beta)},
\]
where $g^*_\alpha$ is the inner product of $T^*_x M$ defined by
\[
g^*_\alpha(\beta, \beta) := \sum_{i,j=1}^n g_{ij}(\alpha) \beta_i \beta_j, \quad \beta = \sum_{i=1}^n \beta_i dx^i.
\]

Proof. Choose local coordinates $(x^i)_{i=1}^n$ around $x$ such that $g_{ij}(v) = \delta_{ij}$ and set

\[
S_x := \left\{ w = \sum_{i=1}^n w^i \frac{\partial}{\partial x^i} \in T_x M \left| \sum_{i=1}^n (w^i)^2 = 1 \right. \right\},
\]
\[
S_x^* := \left\{ \beta = \sum_{i=1}^n \beta_i dx^i \in T_x^* M \left| \sum_{i=1}^n (\beta_i)^2 = 1 \right. \right\}.
\]

First, given $w \in S_x$, we take $\beta \in S_x^*$ such that $\beta(w) = 1$. Then we have $1 = \beta(w) \leq F^*(\beta) F(w)$ and hence
\[
\frac{g_v(w, w)}{F^2(w)} = \frac{1}{F^2(w)} \leq F^*(\beta)^2 = \frac{F^*(\beta)^2}{g^*_\alpha(\beta, \beta)}.
\]

Next, for $\beta' \in S_x^*$, take $w' \in S_x$ with $\beta'(w') = F^*(\beta') F(w')$. Then $F^*(\beta') F(w') = \beta'(w') \leq 1$ and hence $1/F^2(w') \geq F^*(\beta')^2$. This completes the proof. \qed

One can in a similar manner introduce the (2-)uniform convexity constant:

\[
C_F := \sup_{x \in M} \sup_{v, w \in T_x M \setminus 0} \frac{F^2(w)}{g_v(w, w)} = \sup_{x \in M} \sup_{\alpha, \beta \in T_x^* M \setminus 0} \frac{g^*_\alpha(\beta, \beta)}{F^*(\beta)^2} \in [1, \infty].
\tag{2.8}
\]

Again, $C_F = 1$ holds if and only if $(M, F)$ is Riemannian. We remark that $S_F$ and $C_F$ control the reversibility constant $\Lambda_F$ defined in (1.4) as follows.

Lemma 2.4 We have
\[
\Lambda_F \leq \min\{\sqrt{S_F}, \sqrt{C_F}\}.
\]

Proof. For any $v \in TM \setminus 0$, we observe
\[
\frac{F^2(v)}{F^2(-v)} = \frac{g_v(v, v)}{g_v(-v, -v)} = \frac{g_v(-v, -v)}{F^2(-v)} \leq S_F,
\]
and similarly
\[
\frac{F^2(v)}{g_{-v}(v, v)} \geq C_F.
\]
\qed
2.3 Weighted Ricci curvature

The Ricci curvature (as the trace of the flag curvature) on a Finsler manifold is defined by using some connection. Instead of giving a precise definition in coordinates (for which we refer to [BCS]), here we explain a useful interpretation in [Sh, §6.2] going back to (at least) [Au]. Given a unit vector \(v \in T_xM \cap F^{-1}(1)\), we extend it to a \(C^\infty\)-vector field \(V\) on a neighborhood \(U\) of \(x\) in such a way that every integral curve of \(V\) is geodesic, and consider the Riemannian structure \(g_V\) of \(U\) induced from (2.2). Then the Finsler Ricci curvature \(\text{Ric}(v)\) of \(v\) with respect to \(F\) coincides with the Riemannian Ricci curvature of \(v\) with respect to \(g_V\) (in particular, it is independent of the choice of \(V\)).

Inspired by the above interpretation of the Ricci curvature as well as the theory of weighted Ricci curvature (also called the Bakry–Émery–Ricci curvature) of Riemannian manifolds, the weighted Ricci curvature for \((M, F, m)\) was introduced in [Oh2] as follows. Recall that \(m\) is a positive \(C^\infty\)-measure on \(M\), from here on it comes into play.

**Definition 2.5 (Weighted Ricci curvature)** Given a unit vector \(v \in T_xM\), let \(V\) be a \(C^\infty\)-vector field on a neighborhood \(U\) of \(x\) as above. We decompose \(m\) as \(m = e^{-\Psi} \text{vol}_{g_V}\) on \(U\), where \(\Psi \in C^\infty(U)\) and \(\text{vol}_{g_V}\) is the volume form of \(g_V\). Denote by \(\eta : (-\varepsilon, \varepsilon) \rightarrow M\) the geodesic such that \(\dot{\eta}(0) = v\). Then, for \(N \in (-\infty, 0) \cup (n, \infty)\), define

\[
\text{Ric}_N(v) := \text{Ric}(v) + (\Psi \circ \eta)'(0)^2 - \frac{(\Psi \circ \eta)'(0)^2}{N - n}.
\]

We also define as the limits:

\[
\text{Ric}_\infty(v) := \text{Ric}(v) + (\Psi \circ \eta)'(0), \quad \text{Ric}_n(v) := \lim_{N \downarrow n} \text{Ric}_N(v).
\]

For \(c \geq 0\), we set \(\text{Ric}_N(cv) := c^2 \text{Ric}_N(v)\).

We will denote by \(\text{Ric}_N \geq K\), \(K \in \mathbb{R}\), the condition \(\text{Ric}_N(v) \geq KF^2(v)\) for all \(v \in TM\). In the Riemannian case, the study of \(\text{Ric}_\infty\) goes back to Lichnerowicz [Li], he showed a Cheeger–Gromoll type splitting theorem (see [Oh5] for a Finsler counterpart). The range \(N \in (n, \infty)\) has been well studied by Bakry [Ba, §6], Qian [Qi] and many others. The study of the range \(N \in (-\infty, 0)\) is more recent; see [Mi2] for isoperimetric inequalities, [Oh6] for the curvature-dimension condition, and [Wy] for splitting theorems (for \(N \in (-\infty, 1]\)).

It was established in [Oh2] (and [Oh6] for \(N < 0\), [Oh8] for \(N = 0\)) that, for \(K \in \mathbb{R}\), the bound \(\text{Ric}_N \geq K\) is equivalent to Lott, Sturm and Villani’s curvature-dimension condition \(\text{CD}(K, N)\). This extends the corresponding result on weighted Riemannian manifolds and has many geometric and analytic applications (see [Oh2, OS1] among others).

**Remark 2.6 (S-curvature)** For a Riemannian manifold \((M, g, \text{vol}_g)\) endowed with the Riemannian volume measure, clearly we have \(\Psi \equiv 0\) and hence \(\text{Ric}_N = \text{Ric}\) for all \(N\). It is also known that, for Finsler manifolds of Berwald type (i.e., \(\Gamma^k_{ij}\) is constant on each \(T_xM \setminus 0\)), the Busemann–Hausdorff measure satisfies \((\Psi \circ \eta)' \equiv 0\) (in other words, Shen’s S-curvature vanishes, see [Sh, §7.3]). For a general Finsler manifold, however, there may not exist any measure with vanishing S-curvature (see [Oh3] for such an example). This is a reason why we chose to begin with an arbitrary measure \(m\).
For later convenience, we introduce the following notations.

**Definition 2.7 (Reverse Finsler structures)** We define the reverse Finsler structure $\overleftarrow{F}$ of $F$ by $\overleftarrow{F}(v) := F(-v)$.

We will put an arrow $\leftarrow$ on those quantities associated with $\overleftarrow{F}$, we have for example $\overleftarrow{d}(x, y) = d(y, x)$, $\overleftarrow{\text{Ric}}_N(v) = \text{Ric}_N(-v)$ and $\overleftarrow{\nabla} u = -\nabla(-u)$. We say that $(M, F)$ is backward complete if $(M, \overleftarrow{F})$ is forward complete. If $\Lambda_F < \infty$, then these completenesses are mutually equivalent, and we may call it simply completeness.

**2.4 Nonlinear Laplacian and heat flow**

For a differentiable function $u : M \to \mathbb{R}$, the gradient vector at $x$ is defined as the Legendre transform of the derivative of $u$: $\nabla u(x) := \mathcal{L}^*(du(x)) \in T_x M$. If $du(x) \neq 0$, then we can write down in coordinates as

$$\nabla u = \sum_{i,j=1}^{n} g^*_{ij}(du) \frac{\partial u}{\partial x^j} \frac{\partial}{\partial x^i}.$$

We need to be careful when $du(x) = 0$, because $g^*_{ij}(du(x))$ is not defined as well as the Legendre transform $\mathcal{L}^*$ is only continuous at the zero section. Therefore we set

$$M_u := \{ x \in M | du(x) \neq 0 \}.$$

For a twice differentiable function $u : M \to \mathbb{R}$ and $x \in M_u$, we define a kind of Hessian $\nabla^2 u(x) \in T_x^* M \otimes T_x M$ by using the covariant derivative (2.6) as

$$\nabla^2 u(v) := D^u_{\nabla u}(\nabla u)(x) \in T_x M, \quad v \in T_x M.$$

The operator $\nabla^2 u(x)$ is symmetric in the sense that

$$g_{\nabla u}(\nabla^2 u(v), w) = g_{\nabla u}(v, \nabla^2 u(w))$$

for all $v, w \in T_x M$ with $x \in M_u$ (see, for example, [OS3, Lemma 2.3]).

Define the divergence of a differentiable vector field $V$ on $M$ with respect to the measure $\mathbf{m}$ by

$$\text{div}_m V := \sum_{i=1}^{n} \left( \frac{\partial V^i}{\partial x^i} + V^i \frac{\partial \Phi}{\partial x^i} \right), \quad V = \sum_{i=1}^{n} V^i \frac{\partial}{\partial x^i},$$

where we decomposed $\mathbf{m}$ as $d\mathbf{m} = e^\Phi \, dx^1 \, dx^2 \cdots dx^n$. One can rewrite in the weak form as

$$\int_M \phi \text{div}_m V \, d\mathbf{m} = -\int_M d\phi(V) \, d\mathbf{m} \quad \text{for all } \phi \in C_c^\infty(M),$$

that makes sense for measurable vector fields $V$ with $F(V) \in L^1_{\text{loc}}(M)$. Then we define the distributional Laplacian of $u \in H^1_{\text{loc}}(M)$ by $\Delta u := \text{div}_m(\nabla u)$ in the weak sense that

$$\int_M \phi \Delta u \, d\mathbf{m} := -\int_M d\phi(\nabla u) \, d\mathbf{m} \quad \text{for all } \phi \in C_c^\infty(M).$$
Notice that the space $H^1_{\text{loc}}(M)$ is defined solely in terms of the differentiable structure of $M$. Since taking the gradient vector (more precisely, the Legendre transform) is a nonlinear operation, our Laplacian $\Delta$ is a nonlinear operator unless $F$ is Riemannian.

In [OS1, OS3], we have studied the associated nonlinear heat equation $\partial_t u = \Delta u$. In order to recall some results in [OS1], we define the Dirichlet energy of $u \in H^1_{\text{loc}}(M)$ by

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

We remark that $\mathcal{E}(u) < \infty$ does not necessarily imply $\mathcal{E}(-u) < \infty$. Define $H^1_0(M)$ as the closure of $C^\infty_c(M)$ with respect to the (absolutely homogeneous) norm

$$\|u\|_{H^1} := \|u\|_{L^2} + \{\mathcal{E}(u) + \mathcal{E}(-u)\}^{1/2}.$$ 

Note that $(H^1_0(M), \|\cdot\|_{H^1})$ is a Banach space.

**Definition 2.8 (Global solutions)** We say that a function $u$ on $[0, T] \times M$, $T > 0$, is a global solution to the heat equation $\partial_t u = \Delta u$ if it satisfies the following:

1. $u \in L^2([0, T], H^1_0(M)) \cap H^1([0, T], H^{-1}(M))$;
2. For every $\phi \in C^\infty_c(M)$, we have

$$\int_M \phi \cdot \partial_t u_t \, dm = - \int_M d\phi(\nabla u_t) \, dm$$

for almost all $t \in [0, T]$, where we set $u_t := u(t, \cdot)$.

We refer to [Ev] for the notations as in (1). Denoted by $H^{-1}(M)$ is the dual Banach space of $H^1_0(M)$ (so that $H^1_0(M) \subset L^2(M) \subset H^{-1}(M)$). By noticing

$$\int_M |(d\phi - d\tilde{\phi})(\nabla u_t)| \, dm \leq \int_M \max \{F^*(d(\phi - \tilde{\phi})), F^*(d(\tilde{\phi} - \phi))\} F(\nabla u_t) \, dm \leq \{2\mathcal{E}(\phi - \tilde{\phi}) + 2\mathcal{E}(\tilde{\phi} - \phi)\}^{1/2} \cdot \{2\mathcal{E}(u_t)\}^{1/2},$$

the test function $\phi$ can be taken from $H^1_0(M)$. Global solutions can be constructed as gradient curves of the energy functional $\mathcal{E}$ in the Hilbert space $L^2(M)$. We summarize the existence and regularity properties established in [OS1, §§3, 4] in the next theorem.

**Theorem 2.9** Assume $\Lambda_F < \infty$.

(i) For each initial datum $u_0 \in H^1_0(M)$ and $T > 0$, there exists a unique global solution $u$ to the heat equation on $[0, T] \times M$, and the distributional Laplacian $\Delta u_t$ is absolutely continuous with respect to $m$ for all $t \in (0, T)$.

(ii) One can take the continuous version of a global solution $u$, and it enjoys the $H^2_{\text{loc}}$-regularity in $x$ as well as the $C^{1,\alpha}$-regularity for some $\alpha$ in both $t$ and $x$. Moreover, $\partial_t u$ lies in $H^1_{\text{loc}}(M) \cap C(M)$, and further in $H^1_0(M)$ if $S_F < \infty$. 

11
We remark that the usual elliptic regularity yields that \( u \) is \( C^1 \) on \( \bigcup_{t>0}(\{t\} \times M_u) \). The proof of \( \partial_t u \in H^1_0(M) \) under \( S_F < \infty \) can be found in [OS1, Appendix A]. The uniqueness in (i) is a consequence of the convexity of \( F^* \) (see [OS1, Proposition 3.5]).

We finally remark that, by the construction of heat flow as the gradient flow of \( \mathcal{E} \), it is readily seen that:

If \( u_0 \geq 0 \) almost everywhere, then \( u_t \geq 0 \) almost everywhere for all \( t > 0 \). \hspace{1cm} (2.9)

Indeed, if \( u_t < 0 \) on a non-null set, then the curve \( \tilde{u}_t := \max\{u_t, 0\} \) will give a less energy with a less \( L^2 \)-length, a contradiction.

### 2.5 Bochner–Weitzenböck formula

Given \( f \in H^1_{\text{loc}}(M) \) and a measurable vector field \( V \) such that \( V \neq 0 \) almost everywhere on \( M_f = \{ x \in M \mid df(x) \neq 0 \} \), we can define the gradient vector field and the Laplacian on the weighted Riemannian manifold \((M, g_V, m)\) by

\[
\nabla^V f := \begin{cases} 
\sum_{i,j=1}^n g^{ij}(V) \frac{\partial f}{\partial x^j} \frac{\partial}{\partial x^i} & \text{on } M_f, \\
0 & \text{on } M \setminus M_f,
\end{cases} \\
\Delta^V f := \text{div}_m(\nabla^V f),
\]

where the latter is in the sense of distribution. We have \( \nabla^V u = \nabla u \) and \( \Delta^V u = \Delta u \) for \( u \in H^1_{\text{loc}}(M) \) ([OS1, Lemma 2.4]). We also observe that, for \( f_1, f_2 \in H^1_{\text{loc}}(M) \) and \( V \) such that \( V \neq 0 \) almost everywhere,

\[
df_2(\nabla^V f_1) = g_V(\nabla^V f_1, \nabla^V f_2) = df_1(\nabla^V f_2). \hspace{1cm} (2.10)
\]

We established in [OS3, Theorem 3.3] the following key ingredient of the \( \Gamma \)-calculus.

**Theorem 2.10 (Bochner–Weitzenböck formula)** Given \( u \in C^\infty(M) \), we have

\[
\Delta^u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) = \text{Ric}_{\infty}(\nabla u) + \|\nabla^2 u\|_{\text{HS}(\nabla u)}^2 \hspace{1cm} (2.11)
\]

as well as

\[
\Delta^u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) \geq \text{Ric}_N(\nabla u) + \frac{(\Delta u)^2}{N} \hspace{1cm} \text{for } N \in (-\infty, 0) \cup [n, \infty] \text{ point-wise on } M_u,
\]

where \( \|\cdot\|_{\text{HS}(\nabla u)} \) denotes the Hilbert–Schmidt norm with respect to \( g_{\nabla u} \).

In particular, if \( \text{Ric}_N \geq K \), then we have

\[
\Delta^u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) \geq K F^2(\nabla u) + \frac{(\Delta u)^2}{N} \hspace{1cm} (2.12)
\]

on \( M_u \), that we will call the **Bochner inequality**. One can further generalize the Bochner–Weitzenböck formula to a more general class of Hamiltonian systems (by dropping the positive 1-homogeneity; see [Lee, Oh4]).
Theorem 2.13 (Integrated form) Assume $\text{Ric}_N \geq K$ for some $K \in \mathbb{R}$ and $N \in (-\infty, 0) \cup [n, \infty]$. Given $u \in H^2_{\text{loc}}(M) \cap C^1(M)$ such that $\Delta u \in H^1_{\text{loc}}(M)$, we have

$$-\int_M d\phi \left( \nabla^2 u \left[ \frac{f^2}{2} \right] \right) \, dm \geq \int_M \phi \left\{ d(\Delta u)(\nabla u) + KF^2(\nabla u) + \frac{(\Delta u)^2}{N} \right\} \, dm$$

for all bounded nonnegative functions $\phi \in H^1_\varepsilon(M) \cap L^\infty(M)$.

Recall from Theorem 2.9(ii) that global solutions to the heat equation always enjoy $u \in H^2(M) \cap H^2_{\text{loc}}(M) \cap C^1(M)$ and $\Delta u \in H^1_{\text{loc}}(M)$.

3 Linearized semigroups and gradient estimates

In the Bochner–Weitzenböck formula (Theorem 2.10) in the previous section, we used the linearized Laplacian $\nabla^2 u$ induced from the Riemannian structure $g_{\nabla u}$. In the same spirit, we can consider the linearized heat equation associated with a global solution to the heat equation. This technique turned out useful and we have obtained gradient estimates à la Bakry–Émery and Li–Yau in [OS3, §4]. In this section we discuss such a linearization in detail and improve the $L^2$-gradient estimate to an $L^1$-bound (Theorem 3.7).

3.1 Linearized heat semigroups and their adjoints

Let $(u_t)_{t \geq 0}$ be a global solution to the heat equation. We will fix a measurable one-parameter family of non-vanishing vector fields $(V_t)_{t \geq 0}$ such that $V_t = \nabla u_t$ on $M_{u_t}$ for each $t \geq 0$. Given $f \in H^1_0(M)$ and $s \geq 0$, let $(P^s_{s, t}(f))_{t \geq s}$ be the weak solution to the linearized heat equation:

$$\partial_t [P^s_{s, t}(f)] = \Delta^V_p \left[ P^s_{s, t}(f) \right], \quad P^s_{s, t}(f) = f. \quad (3.1)$$

The existence and other properties of the linearized semigroup $P^s_{s, t}$ are summarized in the following proposition.

Proposition 3.1 (Properties of linearized semigroups) Assume that $(M, F, m)$ is complete and satisfies $C_F < \infty$ and $S_F < \infty$, and let $(u_t)_{t \geq 0}$ and $(V_t)_{t \geq 0}$ be as above.
(i) For each \( s \geq 0 \), \( T > 0 \) and \( f \in H_0^1(M) \), there exists a unique weak solution \( \phi_t = P_{s,t}^u(f) \), \( t \in [s,s+T] \), to (3.1). Moreover, \( (\phi_t)_{t \in [s,s+T]} \) lies in \( L^2([s,s+T], H_0^1(M)) \cap H^1([s,s+T], H^{-1}(M)) \) as well as \( C([s,s+T], L^2(M)) \).

(ii) The solution \( (\phi_t)_{t \in [s,s+T]} \) in (i) is Hölder continuous on \((s,s+T) \times M\).

(iii) Assume that either \( m(M) < \infty \) or \( \text{Ric}_\infty \geq K \) for some \( K \in \mathbb{R} \) holds. If \( c \leq f \leq C \) for some \(-\infty < c < C < \infty \), then we have \( c \leq \phi_t \leq C \) almost everywhere for all \( t \in (s,s+T) \).

**Proof.** (i) Let \( s = 0 \) without loss of generality. This unique existence follows from Theorem 4.1 and Remark 4.3 in [LM, Chapter III] (see also [RR, Theorem 11.3]), where \( A(t) \) is assumed to be continuous in \( t \) but it is in fact unnecessary. Precisely, in the notations in [LM], we take \( H = L^2(M), V = H_0^1(M) \), and put \( A_t := -\Delta_t^L : H_0^1(M) \rightarrow H^{-1}(M) \). We deduce with the help of (2.8) that, for any \( h, \tilde{h} \in H_0^1(M) \),

\[
\int_M \tilde{h} \Delta_t^L h \, dm = \int_M g_{\mathcal{L}(V_i)}(dh, d\tilde{h}) \, dm \leq 2\sqrt{\mathcal{E}_t(h)} \sqrt{\mathcal{E}_t(\tilde{h})} \leq 2C_F \sqrt{\mathcal{E}(h)} \sqrt{\mathcal{E}(\tilde{h})}
\]

and

\[
- \int_M h \Delta_t^L \tilde{h} \, dm = 2\mathcal{E}_t(h) \geq \frac{2}{\bar{S}_F} \mathcal{E}(h),
\]

where \( \mathcal{E}_t(h) := (1/2) \int_M g_{\mathcal{L}(V_i)}(dh, d\tilde{h}) \, dm \) denotes the energy functional on \((M, g_{V_i}, m)\). Since \( \Lambda_F < \infty \) by \( C_F < \infty \) (or \( \bar{S}_F < \infty \)), \( \|h\|_{L^2} + \sqrt{\mathcal{E}(h)} \) is comparable with \( \|h\|_{H^1} \). Therefore we have a unique solution \((\phi_t)_{t \in [0,T]} \) to (3.1) with \( f_0 = f \) lying in \( L^2([0,T], H_0^1(M)) \cap H^1([0,T], H^{-1}(M)) \), and also in \( C([0,T], L^2(M)) \) (see [Ev, §5.9.2], [RR, Lemma 11.4]).

(ii) The Hölder continuity is a consequence of the local uniform ellipticity of \( \Delta_t^L \) (see [OS1, Proposition 4.4]).

(iii) This is seen for example by using the fundamental solution \( q(t,x; s,y) \) to the equation \( \partial_t [P_{s,t}^u(f)] = \Delta_t^L [P_{s,t}^u(f)] \) (see [Sal, §6]). We have

\[
\phi_t(x) = \int_M q(t,x; s,y) f(y) \, m(dy),
\]

and \( \int_M q(t,x; s,y) \, m(dy) = 1 \) by \( 1 \in H_0^1(M) \) when \( m(M) < \infty \), or by [Sal, §7] since \( \text{Ric}_\infty \geq K \) implies the squared exponential volume bound as in [St1, Theorem 4.24]). This completes the proof. \( \square \)

The uniqueness in (i) above ensures that \( u_t = P_{s,t}^u(u_s) \). It follows from the non-expansion property,

\[
\frac{d}{dt} \|\phi_t\|_{L^2}^2 = -4\mathcal{E}_t(\phi_t) \leq 0,
\]

that \( P_{s,t}^u \) uniquely extends to a linear contraction semigroup acting on \( L^2(M) \). Notice also that \( f \) is \( C^\infty \) on \( \bigcup_{s<t<s+T} \{t\} \times M_u \).

The operator \( P_{s,t}^u \) is linear but not symmetric (with respect to the \( L^2 \)-inner product). Let us denote by \( \bar{P}_{s,t}^u \) the adjoint operator of \( P_{s,t}^u \). That is to say, given \( \phi \in H_0^1(M) \) and \( t > 0 \), we define \((\bar{P}_{s,t}^u(\phi))_{s \in [0,t]} \) as the solution to the equation

\[
\partial_s [\bar{P}_{s,t}^u(\phi)] = -\Delta_s [\bar{P}_{s,t}^u(\phi)], \quad \bar{P}_{t,t}^u(\phi) = \phi.
\]
Note that
\[\int_M \phi \cdot P_{s,t}^u(f) \, dm = \int_M \hat{P}_{s,t}^u(\phi) \cdot f \, dm \tag{3.3}\]
does indeed hold, since for \( r \in (0, t - s) \)
\[
\partial_r \left[ \int_M \hat{P}_{s+r,t}^u(\phi) \cdot P_{s,s+r}^u(f) \, dm \right]
= -\int_M \Delta^{s+r}[\hat{P}_{s+r,t}^u(\phi)] \cdot P_{s,s+r}^u(f) \, dm + \int_M \hat{P}_{s+r,t}^u(\phi) \cdot \Delta^{s+r}[P_{s,s+r}^u(f)] \, dm
= 0.
\]

One may rewrite (3.2) as
\[
\partial_{\sigma} [\hat{P}_{t-\sigma,t}^u(\phi)] = \Delta^{t-\sigma} [\hat{P}_{t-\sigma,t}^u(\phi)], \quad \sigma \in [0, t],
\]
to see that the adjoint heat semigroup solves the linearized heat equation \textit{backward in time}. (This evolution is sometimes called the \textit{conjugate heat semigroup}, especially in the Ricci flow theory; see for instance [Ch.+ Chapter 5].) Therefore we see in the same way as \( P_{s,t}^u \) that \( \|\hat{P}_{t-\sigma,t}^u(\phi)\|_{L^2} \) is non-increasing in \( \sigma \) and that \( \hat{P}_{t-\sigma,t}^u \) extends to a linear contraction semigroup acting on \( L^2(M) \).

**Remark 3.2** In general, the semigroups \( P_{s,t}^u \) and \( \hat{P}_{s,t}^u \) depend on the choice of an auxiliary vector field \((V_i)_{i \geq 0}\). We will not discuss this issue, but carefully replace \( V_i \) with \( \nabla u_t \) as far as it is possible (with the help of Lemma 2.12).

By a well known technique based on the Bochner inequality (2.12) with \( N = \infty \), we obtained in [OS3, Theorem 4.1] the \textit{\( L^2 \)-gradient estimate} of the following form.

**Theorem 3.3 (\( L^2 \)-gradient estimate, compact case)** Assume that \((M, F, m)\) is compact and satisfies \( \text{Ric}_{\infty} \geq K \) for some \( K \in \mathbb{R} \). Then, given any global solution \((u_t)_{t \geq 0}\) to the heat equation, we have
\[
F^2(\nabla u_t(x)) \leq e^{-2K(t-s)} P_{s,t}^u\left(F^2(\nabla u_s)\right)(x)
\]
for all \( 0 < s < t < \infty \) and \( x \in M \).

We remark that, by Theorem 2.9, \( F^2(\nabla u_s) \in H^1(M) \) and both sides in Theorem 3.3 are Hölder continuous. Let us stress that we use the nonlinear semigroup \((u_s \rightarrow u_t)\) in the LHS, while in the RHS the linearized semigroup \( P_{s,t}^u \) is employed.

**Remark 3.4** In the proof of [OS3, Theorem 4.1], we did not distinguish \( P_{s,t}^u \) and \( \hat{P}_{s,t}^u \) and treated \( P_{s,t}^u \) as a symmetric operator. However, the proof is valid by replacing \( P_{s,t}^u(h) \) with \( \hat{P}_{s,t}^u(h) \). See the proof of Theorem 3.7 below which is based on a similar calculation (with the sharper inequality in Proposition 3.5).
3.2 Improved Bochner inequality

We shall give an inequality improving the Bochner inequality (2.12) with $N = \infty$, that will be used to show the $L^1$-gradient estimate as well as the isoperimetric inequality. In the context of linear diffusion operators, such an inequality can be derived from (2.12) by a self-improvement argument (see [BGL, §C.6], and also [Sav] for an extension to RCD($K, \infty$)-spaces). Here we give a direct proof by calculations in coordinates.

**Proposition 3.5 (Improved Bochner inequality)** Assume $\text{Ric}_\infty \geq K$ for some $K \in \mathbb{R}$. Then we have, for any $u \in C^\infty(M)$,

$$\Delta_u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) - K F^2(\nabla u) \geq d[F(\nabla u)](\nabla^u[F(\nabla u)]) \tag{3.4}$$

point-wise on $M_u$.

**Proof.** By comparing (2.12) with $N = \infty$ and (3.4), it suffices to show

$$4 F^2(\nabla u) \| \nabla^2 u \|_{\text{HS}(\nabla u)} \geq d[F^2(\nabla u)](\nabla^u[F^2(\nabla u)]). \tag{3.5}$$

Fix $x \in M_u$ and choose local coordinates such that $g_{ij}(\nabla u(x)) = \delta_{ij}$. We first calculate the RHS of (3.5) at $x$ as

$$d[F^2(\nabla u)](\nabla^u[F^2(\nabla u)]) = \sum_{i=1}^n \left( \frac{\partial[F^2(\nabla u)]}{\partial x^i} \right)^2$$

$$= \sum_{i=1}^n \left( \frac{\partial}{\partial x^i} \left[ \sum_{j,k=1}^n g_{jk}(du) \frac{\partial u}{\partial x^j} \frac{\partial u}{\partial x^k} \right] \right)^2$$

$$= \sum_{i=1}^n \left( 2 \sum_{j=1}^n \frac{\partial u}{\partial x^j} \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{j,k=1}^n \frac{\partial g_{jk}^*(du)}{\partial x^i} \frac{\partial u}{\partial x^j} \frac{\partial u}{\partial x^k} + \sum_{j,k,l=1}^n \frac{\partial g_{jk}^*(du)}{\partial u_l} \frac{\partial^2 u}{\partial x^i \partial x^j \partial x^k} \right)^2$$

$$= \sum_{i=1}^n \left( 2 \sum_{j=1}^n \frac{\partial u}{\partial x^j} \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{j,k=1}^n \frac{\partial g_{jk}^*(du)}{\partial x^i} \frac{\partial u}{\partial x^j} \frac{\partial u}{\partial x^k} \right)^2,$$

where we used Euler’s theorem (Theorem 2.2, similarly to (2.3)) in the last equality. Next we observe from (2.6) and (2.5) that, again at $x$,

$$\nabla^2 u \left( \frac{\partial}{\partial x^j} \right) = D_{\partial_j}(\nabla u)$$

$$= \sum_{i=1}^n \left\{ \frac{\partial}{\partial x^i} \left[ \sum_{k=1}^n g_{ik}^*(du) \frac{\partial u}{\partial x^k} \right] + \sum_{k=1}^n \Gamma^i_{jk}(\nabla u) \frac{\partial u}{\partial x^k} \right\} \frac{\partial}{\partial x^j}$$

$$= \sum_{i=1}^n \left\{ \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^n \frac{\partial g_{ik}^*(du)}{\partial x^j} \frac{\partial u}{\partial x^k} + \sum_{k=1}^n \gamma^i_{jk}(\nabla u) \frac{\partial u}{\partial x^k} - \sum_{l=1}^n \frac{A_{ijl}(\nabla u)}{F(\nabla u)} \frac{\partial G^l(\nabla u)}{\partial x^j} \right\} \frac{\partial}{\partial x^i}$$

$$= \sum_{i=1}^n \left\{ \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^n \left( \gamma^i_{jk} - \frac{\partial g_{ik}^*(du)}{\partial x^j} \right) \frac{\partial u}{\partial x^k} - \frac{A_{ijl}(\nabla u)}{F(\nabla u)} \frac{\partial G^l(\nabla u)}{\partial x^j} \right\} \frac{\partial}{\partial x^i}.$$
In the last line we used
\[ \frac{\partial g_{ik}}{\partial x^j}(du(x)) = -\frac{\partial g_{ik}}{\partial x^j}(\nabla u(x)). \]

Hence we deduce from the Cauchy–Schwarz inequality, (2.3) and (2.4) that

\[ F^2(\nabla u)\|\nabla^2 u\|_{HS(\nabla u)}^2 \]
\[ = \sum_{j=1}^{n} \left( \frac{\partial u}{\partial x^j} \right)^2 \cdot \sum_{i,j=1}^{n} \left( \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^{n} \left( \gamma_{jk} - \frac{\partial g_{ik}}{\partial x^j} (\nabla u) \frac{\partial u}{\partial x^k} - \sum_{k=1}^{n} A_{ijk} G_k^k (\nabla u) \right)^2 \right) \]
\[ \geq \sum_{i=1}^{n} \left( \frac{\partial u}{\partial x^i} \right)^2 \left( \frac{\partial u}{\partial x^i} \right)^2 + \sum_{j,k=1}^{n} \left( \gamma_{jk} - \frac{\partial g_{ik}}{\partial x^j} (\nabla u) \frac{\partial u}{\partial x^k} - \sum_{k=1}^{n} A_{ijk} G_k^k (\nabla u) \right)^2 \]
\[ = \sum_{i=1}^{n} \left( \frac{\partial u}{\partial x^i} \right)^2 \left( \frac{\partial u}{\partial x^i} \right)^2 - \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial g_{ik}}{\partial x^j} (\nabla u) \frac{\partial u}{\partial x^j} \frac{\partial u}{\partial x^k} \right)^2. \]

This completes the proof of (3.5) as well as (3.4). □

The following integrated form can be shown in the same way as Theorem 2.13, we refer to [OS3, Theorem 3.6] for details.

**Corollary 3.6 (Integrated form)** Assume \( \text{Ric}_\infty \geq K \) for some \( K \in \mathbb{R} \). Given \( u \in H^2_{\text{loc}}(M) \cap C^1(M) \) such that \( \Delta u \in H^1_{\text{loc}}(M) \), we have

\[ - \int_M d\phi \left( \nabla u \left[ \frac{F^2(\nabla u)}{2} \right] \right) \ dm \]
\[ \geq \int_M \phi \left\{ d(\Delta u)(\nabla u) + K F^2(\nabla u) + d[F(\nabla u)](\nabla^2 u[F(\nabla u)]) \right\} \ dm \]

for all bounded nonnegative functions \( \phi \in H^1_c(M) \cap L^\infty(M) \).

### 3.3 \( L^1 \)-gradient estimate

The improved Bochner inequality (3.4) yields the following \( L^1 \)-gradient estimate, under a technical (likely redundant) assumption that \( d[F(\nabla u_t)](\nabla^2 u_t[F(\nabla u_t)]) \in L^1(M) \) for all \( t > 0 \), which holds in the compact case thanks to the \( H^2_{\text{loc}} \)-regularity (recall Theorem 2.9).

**Theorem 3.7 (\( L^1 \)-gradient estimate)** Let \( (M, F, m) \) be complete and satisfy \( \text{Ric}_\infty \geq K \), \( C_F < \infty \) and \( S_F < \infty \), and \( (u_t)_{t \geq 0} \) be a global solution to the heat equation with \( u_0 \in C_c(M) \). We further assume that

\[ d[F(\nabla u_t)](\nabla^2 u_t[F(\nabla u_t)]) \in L^1(M) \] (3.6)

for all \( t > 0 \). Then we have

\[ F(\nabla u_t(x)) \leq e^{-K(t-s)} P_{s,t}^\nabla u(F(\nabla u_s))(x) \]

for all \( 0 \leq s < t < \infty \) and \( x \in M \).
Proof. Notice first that \( F(\nabla u_0) \in H_0^1(M) \cap L^\infty(M) \) since \( u_0 \in C_0^\infty(M) \). Fix arbitrary \( \varepsilon > 0 \) and let us consider the function

\[
\xi_\sigma := \sqrt{e^{-2K\sigma}F^2(\nabla u_{t-\sigma}) + \varepsilon}, \quad 0 \leq \sigma \leq t - s.
\]

Note from the proof of [OS3, Theorem 4.1] that

\[
\frac{\partial}{\partial \sigma} \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] = -\frac{\partial}{\partial t} \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] = -d(\Delta u_{t-\sigma})(\nabla u_{t-\sigma}). \quad (3.7)
\]

Hence we have, on one hand,

\[
\partial_\sigma \xi_\sigma = -\frac{e^{-2K\sigma}}{\xi_\sigma} \left\{KF^2(\nabla u_{t-\sigma}) + d(\Delta u_{t-\sigma})(\nabla u_{t-\sigma}) \right\}.
\]

On the other hand, for any nonnegative function \( \phi \in C_c^\infty(M) \), we observe

\[
\int_M d\phi(\nabla \nabla_{u_{t-\sigma}} \xi_\sigma) \, dm = \int_M \frac{e^{-2K\sigma}}{\xi_\sigma} d\phi \left( \nabla \nabla_{u_{t-\sigma}} \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] \right) \, dm
\]

\[
= \int_M \left( d \left( \frac{e^{-2K\sigma}}{\xi_\sigma} \right) + \frac{e^{-2K\sigma}}{\xi_\sigma^2} d\xi_\sigma \right) \left( \nabla \nabla_{u_{t-\sigma}} \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] \right) \, dm
\]

\[
+ \int_M \frac{e^{-4K\sigma}}{\xi_\sigma} d \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] \left( \nabla \nabla_{u_{t-\sigma}} \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] \right) \, dm
\]

\[
\leq \int_M \frac{e^{-2K\sigma}}{\xi_\sigma} \left( \nabla \nabla_{u_{t-\sigma}} \left[ \frac{F^2(\nabla u_{t-\sigma})}{2} \right] \right) \, dm
\]

\[
+ \int_M \frac{e^{-4K\sigma}}{\xi_\sigma} d[F(\nabla u_{t-\sigma})] \left( \nabla \nabla_{u_{t-\sigma}} [F(\nabla u_{t-\sigma})] \right) \, dm,
\]

where we used \( F^2(\nabla u_{t-\sigma}) \leq e^{2K\sigma}\xi_\sigma^2 \) in the last inequality. Therefore the improved Bochner inequality (Corollary 3.6) shows that

\[
\Delta \nabla_{u_{t-\sigma}} \xi_\sigma + \partial_\sigma \xi_\sigma \geq 0 \quad (3.8)
\]

in the weak sense. Notice that the test function \( \phi \) can be in fact taken from \( H_0^1(M) \cap L^\infty(M) \) thanks to the hypothesis (3.6) and \( \xi_\sigma \geq \sqrt{\varepsilon} \).

For a nonnegative function \( \varphi \in C_c^\infty(M) \) and \( \sigma \in (0, t - s) \), set

\[
\Phi(\sigma) := \int_M \varphi \cdot F_{t-\sigma}(\xi_\sigma) \, dm = \int_M \hat{P}_{t-\sigma}(\varphi) \cdot \xi_\sigma \, dm.
\]

We deduce from (3.2) and (2.10) that

\[
\Phi'(\sigma) = \int_M \hat{P}_{t-\sigma}(\varphi) \cdot \partial_\sigma \xi_\sigma \, dm - \int_M d\xi_\sigma \left( \nabla \nabla_{u_{t-\sigma}} \left[ \hat{P}_{t-\sigma}(\varphi) \right] \right) \, dm
\]

\[
= \int_M \hat{P}_{t-\sigma}(\varphi) \cdot \partial_\sigma \xi_\sigma \, dm - \int_M d[\hat{P}_{t-\sigma}(\varphi)] \left( \nabla \nabla_{u_{t-\sigma}} \xi_\sigma \right) \, dm.
\]

18
Therefore we can apply (3.8) with the test function $\hat{P}_t^{\nabla u}(\varphi)$ (thanks to Proposition 3.1) to obtain $\Phi'(\sigma) \geq 0$. This implies
\[
\int_M \varphi \cdot \xi_0 \, dm \leq \int_M \varphi \cdot P_{s,t}^{\nabla u}(\xi_{t-s}) \, dm.
\]
By the arbitrariness of $\varphi$ and $\varepsilon$, we have
\[
F(\nabla u_t) \leq e^{-K(t-s)} P_{s,t}^{\nabla u}(F(\nabla u_s))
\]
almost everywhere. Since both sides are Hölder continuous (Proposition 3.1(ii)), this completes the proof.

It is a standard fact that the $L^1$-gradient estimate implies the $L^2$-bound.

**Corollary 3.8 (L²-gradient estimate, noncompact case)** Let $(M, F, m)$ be complete and satisfy $\text{Ric} \geq K$, $C_F < +\infty$ and $S_F < +\infty$, and $(u_t)_{t \geq 0}$ be a global solution to the heat equation with $u_0 \in C_0^\infty(M)$ and satisfying (3.6) for all $t > 0$. Then we have
\[
F^2(\nabla u_t(x)) \leq e^{-2K(t-s)} P_{s,t}^{\nabla u}(F^2(\nabla u_s))(x)
\]
for all $0 \leq s < t < \infty$ and $x \in M$.

**Proof.** This is a consequence of a kind of Jensen’s inequality:
\[
P_{s,t}^{\nabla u}(f)^2 \leq P_{s,t}^{\nabla u}(f^2)
\]
for $f \in L^2(M) \cap L^\infty(M)$. For $\psi \in C_0^\infty(M)$ with $0 \leq \psi \leq 1$ and $r \in \mathbb{R}$, we have
\[
0 \leq P_{s,t}^{\nabla u}(rf + \psi)^2 = r^2 P_{s,t}^{\nabla u}(f^2) + 2r P_{s,t}^{\nabla u}(f \psi) + P_{s,t}^{\nabla u}(\psi^2)
\]
\[
\leq r^2 P_{s,t}^{\nabla u}(f^2) + 2r P_{s,t}^{\nabla u}(f \psi) + 1.
\]
Letting $f \psi \to f$ in $L^2(M)$, we find $r^2 P_{s,t}^{\nabla u}(f^2) + 2r P_{s,t}^{\nabla u}(f) + 1 \geq 0$ for all $r \in \mathbb{R}$. Hence $P_{s,t}^{\nabla u}(f)^2 - P_{s,t}^{\nabla u}(f^2) \leq 0$ as desired. \qed

### 3.4 On the hypothesis (3.6)

The hypothesis (3.6) seems redundant and indeed unnecessary for weighted Riemannian manifolds and RCD-spaces. Especially, when $K > 0$, the Gaussian decay of the measure ([St1, Theorem 4.26]) could imply (3.6). Let us give some more comments on (3.6).

#### 3.4.1 Weighted Riemannian case

We essentially followed the proof of [BGL, Theorem 3.2.4] in Theorem 3.7. Then we have
\[
d\xi_\sigma(\nabla u_{t-\sigma}) \leq e^{-2K\sigma} d[F(\nabla u_{t-\sigma})](\nabla u_{t-\sigma}) [F(\nabla u_{t-\sigma})],
\]
and the improved Bochner inequality (Proposition 3.5) implies
\[
\int_M d[F(\nabla u)][(\nabla u F(\nabla u))] \, dm \leq \|\Delta u\|_{L^2}^2 + |K| \cdot \|u\|_{L^2} \|\Delta u\|_{L^2}
\]

19
for \( u \in C_c^\infty(M) \). Now in [BGL], for a linear operator \( L \), we make use of the density of \( A_0 = C_c^\infty(M) \) in the domain \( D(L) \) with respect to the norm

\[
\| f \|_{D(L)} := \left( \| f \|^2_{L^2} + \| Lf \|^2_{L^2} \right)^{1/2}
\]
to extend the above estimate to \( D(L) \). This density is a consequence of the hypo-ellipticity (see [BGL, Proposition 3.2.1]), which is defined by the property that any solution to \( L^*f = \lambda f \) is smooth (see also [BGL, Definition 3.3.8], typically \( A = C^\infty(M) \)). This is not the case for operators with nonsmooth coefficients, thereby it is unclear if we can apply this method in the Finsler case (to the linearized Laplacian \( \Delta\nabla u \)).

### 3.4.2 RCD-case

In RCD\((K, \infty)\)-spaces, we obtain the Wasserstein contraction estimate of heat flow by the convexity of the relative entropy, and then the gradient estimates follow by the duality argument. Moreover, we can obtain the Bochner inequality by differentiating the gradient estimate (see [AGS1, AGS2, GKO, Sav] for details).

This method could avoid the use of the functional analytic argument involving \( A_0 \) and \( A \), and what is important and interesting here is that the Bochner inequality derived from the gradient estimate is of the form:

\[
\int \Delta \phi \cdot \frac{\nabla u^2}{2} \, dm \geq \int \phi \left\{ d(\Delta u)(\nabla u) + K|\nabla u|^2 \right\} \, dm
\]

for \( u \in D(\Delta) \) with \( \Delta u \in H^1 \) and \( \phi \in D(\Delta) \cap L^\infty \) with \( \Delta \phi \in L^\infty \). In the LHS, what we have directly from the point-wise Bochner inequality is

\[
\int \phi \cdot \Delta \left[ \frac{\nabla u^2}{2} \right] \, dm,
\]

and modifying this into the above LHS requires an approximation of \( \phi \) by functions \( \phi_k \) in \( C_c^\infty \) such that \( \Delta \phi_k \to \Delta \phi \), namely the density of \( C_c^\infty \) in the \( D(\Delta) \)-norm as in the approach of [BGL].

In the Finsler case, we know that the Wasserstein contraction fails (see Remark 3.10 below). Nonetheless, if one can show the Bochner inequality in the above form as well as \( F(\nabla u) \in L^\infty(M) \), then it follows from the argument along [Sav, Lemma 3.2] that \( d[F(\nabla u)](\nabla \nabla F(\nabla u)) \in L^1(M) \) and we obtain the gradient estimates.

### 3.5 Characterizations of lower Ricci curvature bounds

We close the section with several characterizations of the lower Ricci curvature bound \( \text{Ric}_\infty \geq K \).

**Theorem 3.9 (Characterizations of \( \text{Ric}_\infty \geq K \))** Let \((M, F, m)\) be complete and satisfy \( C_F < \infty \) and \( S_F < \infty \). We assume that (3.6) holds for all solutions \((u_t)_{t \geq 0}\) to the heat equation with \( u_0 \in C_c^\infty(M) \). Then, for each \( K \in \mathbb{R} \), the following are equivalent:

1. \( \text{Ric}_\infty \geq K \).
(II) The Bochner inequality
\[ \Delta \nabla u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) \geq K F^2(\nabla u) \]
holds on \( M_u \) for all \( u \in C^\infty(M) \).

(III) The improved Bochner inequality
\[ \Delta \nabla u \left[ \frac{F^2(\nabla u)}{2} \right] - d(\Delta u)(\nabla u) - K F^2(\nabla u) \geq d(F(\nabla u))(\nabla u[F(\nabla u)]) \]
holds on \( M_u \) for all \( u \in C^\infty(M) \).

(IV) The \( L^2 \)-gradient estimate
\[ F^2(\nabla u) \leq e^{-2K(t-s)} P_{s,t}^u \left( F^2(\nabla u_s) \right), \quad 0 \leq s < t < \infty, \]
holds for all global solutions \( (u_t)_{t \geq 0} \) to the heat equation with \( u_0 \in C^\infty_c(M) \).

(V) The \( L^1 \)-gradient estimate
\[ F(\nabla u) \leq e^{-K(t-s)} P_{s,t}^u \left( F(\nabla u_s) \right), \quad 0 \leq s < t < \infty, \]
holds for all global solutions \( (u_t)_{t \geq 0} \) to the heat equation with \( u_0 \in C^\infty_c(M) \).

Proof. We have shown (I) \( \Rightarrow \) (III) in Proposition 3.5, (III) \( \Rightarrow \) (V) in Theorem 3.7, and (V) \( \Rightarrow \) (IV) in Corollary 3.8. One can deduce (IV) \( \Rightarrow \) (II) from the proof of [OS3, Theorem 4.1] or by differentiating \( F^2(\nabla u) \leq e^{-2K_t} P_{0,t}^u \left( F^2(\nabla u_0) \right) \) at \( t = 0 \) (recall (3.7), see also [GKO]). Let us finally prove (II) \( \Rightarrow \) (I). Given \( v_0 \in T_{x_0} M \setminus 0 \), fix local coordinates \( (x^i)_{i=1}^n \) around \( x_0 \) with \( g_{ij}(v_0) = \delta_{ij} \) and \( x^i(x_0) = 0 \) for all \( i \). Consider the function
\[ u := \sum_{i=1}^n v_0^i x^i + \frac{1}{2} \sum_{i,j,k=1}^n \Gamma^k_{ij}(v_0) v_0^k x^i x^j \]
on a neighborhood of \( x_0 \), and observe that \( \nabla u(x_0) = v_0 \) as well as \( (\nabla^2 u)|_{T_{x_0}M} = 0 \) (see [OS3, Lemma 2.3] for the precise expression in coordinates of \( \nabla^2 u \)). Then the Bochner–Weitzenböck formula (2.11) and (II) imply
\[ \operatorname{Ric}_{\infty}(v_0) = \Delta \nabla u \left[ \frac{F^2(\nabla u)}{2} \right](x_0) - d(\Delta u)(\nabla u)(x_0) \geq K F^2(v_0). \]
This completes the proof. \( \square \)

Remark 3.10 (The lack of contraction) In the Riemannian context, lower Ricci curvature bounds are also equivalent to contraction estimates of heat flow with respect to the Wasserstein distance (we refer to [vRS] for the Riemannian case, and [EKS] for the case of RCD-spaces). More generally, for linear semigroups, gradient estimates are directly equivalent to the corresponding contraction properties (see [Ku]). In our Finsler setting, however, the lack of the commutativity (see [OP]) prevents such a contraction estimate, at least in the same form (see [OS2] for details).
Remark 3.11 (Similarities to (super) Ricci flow theory) The methods in this section have connections with the Ricci flow theory. Ricci flow provides time-dependent Riemannian metrics obeying a kind of heat equation on the space of Riemannian metrics, while we considered the time-dependent (singular) Riemannian structures $g_{\nabla u}$ for $u$ solving the heat equation. More precisely, what corresponds to our lower Ricci curvature bound is super Ricci flow (super-solutions to the Ricci flow equation). We refer to [MT] for an inspiring work on a characterization of super Ricci flow in terms of the contraction of heat flow, and to [St3] for a recent investigation of super Ricci flow on time-dependent metric measure spaces including various characterizations related to Theorem 3.9. Then, again, what is missing in our Finsler setting is the contraction property, for which the Riemannian nature of the space is necessary.

4 Bakry–Ledoux’s isoperimetric inequality

This section is devoted to the isoperimetric inequality, as a geometric application of the improved Bochner inequality (Proposition 3.5). We will assume $\text{Ric}_\infty \geq K > 0$, then $m(M) < \infty$ holds (see [St1, Theorem 4.26]) and hence we can normalize $m$ as $m(M) = 1$ without changing $\text{Ric}_\infty$ (cm with $c > 0$ gives the same weighted Ricci curvature as $m$).

For a Borel set $A \subset M$, define the Minkowski exterior boundary measure as

$$m^+(A) := \liminf_{\varepsilon \downarrow 0} \frac{m(B^+(A, \varepsilon)) - m(A)}{\varepsilon},$$

where $B^+(A, \varepsilon) := \{y \in M \mid \inf_{x \in A} d(x, y) < \varepsilon\}$ is the forward $\varepsilon$-neighborhood of $A$. Then the (forward) isoperimetric profile $I_{(M, F, m)} : [0, 1] \rightarrow [0, \infty)$ of $(M, F, m)$ is defined by

$$I_{(M, F, m)}(\theta) := \inf\{m^+(A) \mid A \subset M : \text{Borel set with } m(A) = \theta\}.$$

Clearly $I_{(M, F, m)}(0) = I_{(M, F, m)}(1) = 0$. The following is our main result (stated as Theorem in the introduction).

Theorem 4.1 (Bakry–Ledoux’s isoperimetric inequality) Let $(M, F)$ be complete and satisfy $\text{Ric}_\infty \geq K > 0$, $m(M) = 1$, $C_F < \infty$ and $S_F < \infty$. We assume that (3.6) holds for all solutions $(u_t)_{t \geq 0}$ to the heat equation with $u_0 \in C_c^\infty(M)$. Then we have

$$I_{(M, F, m)}(\theta) \geq I_K(\theta)$$

for all $\theta \in [0, 1]$, where

$$I_K(\theta) := \sqrt{\frac{K}{2\pi}} e^{-K \theta^2/2}$$

with $\theta = \int_{-\infty}^{c(\theta)} \sqrt{\frac{K}{2\pi}} e^{-K a^2/2} da$.

Recall that, under $C_F < \infty$ or $S_F < \infty$, the forward completeness is equivalent to the backward completeness by Lemma 2.4. In the Riemannian case, the inequality (4.1) is due to Bakry and Ledoux [BL] (see also [BGL, §8.5.2]) and can be regarded as the dimension-free version of Lévy–Gromov’s isoperimetric inequality (see [Lé1, Lé2, Gr]). Lévy–Gromov’s classical isoperimetric inequality asserts that the isoperimetric profile of
an $n$-dimensional Riemannian manifold $(M, g)$ with $\text{Ric} \geq n - 1$ is bounded below by the profile of the unit sphere $\mathbb{S}^n$ (both spaces are equipped with the normalized volume measures). In (4.1), the role of the unit sphere is played by the real line $\mathbb{R}$ equipped with the Gaussian measure $\sqrt{K/2\pi} e^{-K x^2/2} dx$, thereby (4.1) is also called the Gaussian isoperimetric inequality.

In [Oh8], generalizing Cavalletti and Mondino’s localization technique in [CM], we showed the slightly weaker inequality (recall the introduction)

$$I_{(M, F, m)}(\theta) \geq \Lambda_F^{-1} \cdot I_K(\theta)$$

under the finite reversibility $\Lambda_F < \infty$ (but without $C_F < \infty$ nor $S_F < \infty$). In fact we have treated in [Oh8] the general curvature-dimension-diameter bound $\text{Ric}_N \geq K$ and $\text{diam } M \leq D$ (in accordance with [Mi1]). Theorem 4.1 sharpens the estimate in [Oh8] in the special case of $N = D = \infty$ and $K > 0$.

### 4.1 Ergodicity

We begin with some properties induced from our hypothesis $\text{Ric}_\infty \geq K > 0$.

**Lemma 4.2 (Global Poincaré inequality)** Suppose that $(M, F, m)$ is forward or backward complete, $\text{Ric}_\infty \geq K > 0$ and $m(M) = 1$. Then we have, for any locally Lipschitz function $f \in H^1_0(M)$,

$$\int_M f^2 \, dm - \left( \int_M f \, dm \right)^2 \leq \frac{1}{K} \int_M F^\ast (df)^2 \, dm. \quad (4.2)$$

**Proof.** It is well known that the curvature bound $\text{Ric}_\infty \geq K$ (or $\text{CD}(K, \infty)$) implies the log-Sobolev inequality,

$$\int_M \rho \log \rho \, dm \leq \frac{1}{2K} \int_M \frac{F^\ast (d\rho)^2}{\rho} \, dm \quad (4.3)$$

for nonnegative locally Lipschitz functions $\rho$ with $\int_M \rho \, dm = 1$, and that (4.2) follows from (4.3) (see [OV, LV, Vi, Oh2]). Here we explain the latter step for thoroughness.

By truncation, let us assume that $f$ is bounded. Since

$$\int_M f^2 \, dm - \left( \int_M f \, dm \right)^2 = \int_M \left( f - \int_M f \, dm \right)^2 \, dm,$$

we can further assume that $\int_M f \, dm = 0$. There is nothing to prove if $f \equiv 0$, thereby assume $\|f\|_{L^\infty} > 0$. For $\varepsilon \in \mathbb{R}$ with $|\varepsilon| < \|f\|_{L^\infty}$, we consider the probability measure $(1 + \varepsilon f) m$. Then the log-Sobolev inequality for $\rho_\varepsilon := 1 + \varepsilon f$ under $\text{Ric}_\infty \geq K$ implies

$$\int_M (1 + \varepsilon f) \log(1 + \varepsilon f) \, dm \leq \frac{1}{2K} \int_M \frac{\varepsilon^2 F^\ast (df)^2}{1 + \varepsilon f} \, dm.$$
where $O(\varepsilon^3)$ in the LHS is uniform in $M$ thanks to the boundedness of $f$. Hence we have

$$
\frac{\varepsilon^2}{2} \int_M f^2 \, dm \leq \frac{1}{1 - \varepsilon \|f\|_{L^\infty}} \frac{\varepsilon^2}{2K} \int_M F^*(df)^2 \, dm + O(\varepsilon^3).
$$

Dividing both sides by $\varepsilon^2$ and letting $\varepsilon \to 0$ implies (4.2). □

The LHS of (4.2) is the \textit{variance} of $f$:

$$
\text{Var}_m(f) := \int_M f^2 \, dm - \left( \int_M f \, dm \right)^2.
$$

We next show that the Poincaré inequality (4.2) yields the exponential decay of the variance and a kind of \textit{ergodicity} along heat flow (similarly to [BGL, §4.2]), which is one of the key ingredients in the proof of Theorem 4.1 (see the proof of Corollary 4.5). Given a global solution $(u_t)_{t \geq 0}$ to the heat equation, since the finiteness of the total mass together with $\Lambda_F < \infty$ and the completeness implies $1 \in H^1_0(M)$, we observe the mass conservation:

$$
\int_M P^\nabla_{s,t} f \, dm = \int_M f \, dm \quad (4.4)
$$

for any $f \in H^1_0(M)$ and $0 \leq s < t < \infty$.

**Proposition 4.3 (Variance decay and ergodicity)** Assume that $(M, F, m)$ is complete and satisfies $C_F < \infty$, $S_F < \infty$, $\text{Ric}_\infty \geq K > 0$ and $m(M) = 1$. Then we have, given any global solution $(u_t)_{t \geq 0}$ to the heat equation and $f \in H^1_0(M)$,

$$
\text{Var}_m\big(P^\nabla_{s,t} (f)\big) \leq e^{-2K(t-s)/S_F} \text{Var}_m(f)
$$

for all $0 \leq s < t < \infty$. In particular, $P^\nabla_{s,t} f$ converges to the constant function $\int_M f \, dm$ in $L^2(M)$ as $t \to \infty$.

**Proof.** Put $f_t := P^\nabla_{s,t} f$, then $\int_M f_t \, dm = \int_M f \, dm$ holds by (4.4). It follows from Lemmas 2.3, 4.2 that

$$
\frac{d}{dt} \left[ \text{Var}_m(f_t) \right] = -2 \int_M df_t (\nabla V_t f_t) \, dm = -2 \int_M g^*_L(V_t)(df_t, df_t) \, dm \\
\leq -\frac{2}{S_F} \int_M F^*(df_t)^2 \, dm \leq -\frac{2K}{S_F} \text{Var}_m(f_t).
$$

Hence $e^{2Kt/S_F} \text{Var}_m(f_t)$ is non-increasing in $t$, this completes the proof of the first assertion. Then the second assertion is straightforward since

$$
\text{Var}_m(f_t) = \int_M \left( f_t - \int_M f \, dm \right)^2 \, dm \to 0 \quad (t \to \infty).
$$

□
4.2 Key estimate

We next prove a key estimate which would have further applications (see [BL]). Define
\[
\varphi(c) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} e^{-b^2/2} \, db, \quad c \in \mathbb{R},
\]
\[
\mathcal{N}(\theta) := \varphi'(\mathcal{N}(\theta)) = \frac{e^{-\mathcal{N}(\theta)/2}}{\sqrt{2\pi}}, \quad \theta \in (0, 1).
\]

We set also \(\mathcal{N}(0) = \mathcal{N}(1) = 0\). Observe that \(\mathcal{N}' = -\varphi^{-1}\) and \(\mathcal{N}'' = -1/\mathcal{N}\) on \((0, 1)\).

**Theorem 4.4** Assume that \((M, F, m)\) is complete and satisfies \(\text{Ric}_{\infty} \geq K\) for some \(K \in \mathbb{R}, C_F < \infty, S_F < \infty\) and \(m(M) < \infty\). Then, given a global solution \((u_t)_{t \geq 0}\) to the heat equation with \(u_0 \in C_c^\infty(M)\), \(0 \leq u_0 \leq 1\) and satisfying (3.6), we have
\[
\sqrt{\mathcal{N}^2(u_t) + \alpha F^2(\nabla u_t)} \leq P_{0,t}^{\nabla u}\left(\sqrt{\mathcal{N}^2(u_0) + c_\alpha(t)F^2(\nabla u_0)}\right)
\]
on \(M\) for all \(\alpha \geq 0\) and \(t > 0\), where
\[
c_\alpha(t) := \frac{1 - e^{-2Kt}}{K} + \alpha e^{-2Kt} > 0
\]
and \(c_\alpha(t) := 2t + \alpha\) when \(K = 0\).

For simplicity, we suppressed the dependence of \(c_\alpha\) on \(K\).

**Proof.** By replacing \(u_0\) with \((1 - 2\varepsilon)u_0 + \varepsilon\), we can assume \(\varepsilon \leq u_0 \leq 1 - \varepsilon\) for some \(\varepsilon > 0\), and then we have \(\varepsilon \leq u_0 \leq 1 - \varepsilon\) for all \(t > 0\) (recall (2.9)). Fix \(t > 0\) and put
\[
\zeta_s := \sqrt{\mathcal{N}^2(u_s) + c_\alpha(t-s)F^2(\nabla u_s)}, \quad 0 \leq s \leq t
\]
(compare this function with \(\xi_s\) in the proof of Theorem 3.7). Then (4.5) is written as \(\zeta_s \leq P_{0,t}^{\nabla u}(\zeta_0)\) and it suffices to show \(\partial_s[P_{s,t}^{\nabla u}(\zeta_s)] \leq 0\) in the weak sense. Observe from (3.3) and (3.2) that, for any nonnegative \(\phi \in C_c^\infty((0, t) \times M)\),
\[
\int_0^t \int_M \partial_s \phi_s \cdot P_{s,t}^{\nabla u}(\zeta_s) \, d\mu \, ds = \int_0^t \int_M \widehat{P}_{t-s}^{\nabla u}(\partial_s \phi_s) \cdot \zeta_s \, d\mu \, ds
\]
\[
= \int_0^t \int_M \left\{ \partial_s[\widehat{P}_{s,t}^{\nabla u}(\phi_s)] + \Delta^s[\widehat{P}_{s,t}^{\nabla u}(\phi_s)] \right\} \cdot \zeta_s \, d\mu \, ds
\]
\[
= \int_0^t \int_M \widehat{P}_{s,t}^{\nabla u}(\phi_s) \cdot (\Delta^{s,t} \zeta_s - \partial_s \zeta_s) \, d\mu \, ds, \quad (4.6)
\]
where in the second equality we deduce from the linearity of \(\widehat{P}_{s,t}^{\nabla u}\) that
\[
\int_M \partial_s[\widehat{P}_{s,t}^{\nabla u}(\phi_s)] \cdot \zeta_s \, d\mu
\]
\[
= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_M \{ \widehat{P}_{s,t}^{\nabla u}(\phi_{s+\varepsilon}) - \widehat{P}_{s,t}^{\nabla u}(\phi_{s-\varepsilon}) \} \cdot \zeta_s \, d\mu + \int_M \widehat{P}_{s,t}^{\nabla u}(\partial_s \phi_s) \cdot \zeta_s \, d\mu
\]
\[
= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^{s+\varepsilon} \int_M \zeta_s \cdot (\nabla \nabla u_s + 2 \widehat{P}_{s,t}^{\nabla u}(\phi_{s+\varepsilon})) \, d\mu \, ds + \int_M \widehat{P}_{s,t}^{\nabla u}(\partial_s \phi_s) \cdot \zeta_s \, d\mu
\]
\[
= \int_M d\zeta_s (\nabla \nabla u_s \cdot \widehat{P}_{s,t}^{\nabla u}(\phi_s)) \, d\mu + \int_M \widehat{P}_{s,t}^{\nabla u}(\partial_s \phi_s) \cdot \zeta_s \, d\mu
\]

25
for almost every $s$. We shall show that the RHS of (4.6) is nonnegative.

We first calculate by using (3.7) and $c_\alpha = 2(1 - Kc_\alpha)$ as

$$
\partial_s \zeta_s = \frac{1}{\zeta_s} \left\{ \mathcal{N}'(u_s) \mathcal{N}'(u_s) \mathcal{N}u_s + (Kc_\alpha(t - s) - 1) F^2(\nabla u_s) + c_\alpha(t - s) d(\Delta u_s)(\nabla u_s) \right\}.
$$

Next, we have

$$
\nabla \nabla^s \zeta_s = \frac{1}{\zeta_s} \left\{ \mathcal{N}'(u_s) \mathcal{N}'(u_s) \nabla u_s + \frac{c_\alpha(t - s)}{2} \nabla \nabla^s u_s [F^2(\nabla u_s)] \right\}.
$$

Hence

$$
\Delta \nabla^s \zeta_s = \frac{\mathcal{N}'(u_s) \mathcal{N}'(u_s)}{\zeta_s} \Delta u_s + \frac{\mathcal{N}'(u_s)^2 - 1}{\zeta_s} F^2(\nabla u_s) - \frac{\mathcal{N}'(u_s)}{\zeta_s} d\zeta_s(\nabla u_s)
$$

$$
+ \frac{c_\alpha(t - s)}{2\zeta_s} \Delta \nabla^s u_s [F^2(\nabla u_s)] - \frac{c_\alpha(t - s)}{2\zeta_s^2} d\zeta_s(\nabla \nabla^s u_s [F^2(\nabla u_s)]),
$$

where we used $\mathcal{N}'' = -1/\mathcal{N}$ and $\Delta \nabla^s u_s [F^2(\nabla u_s)]$ is understood in the weak sense.

Now we apply the improved Bochner inequality (Corollary 3.6) to obtain

$$
\Delta \nabla^s \zeta_s - \partial_s \zeta_s = \frac{\mathcal{N}'(u_s)^2 - Kc_\alpha(t - s)}{\zeta_s} F^2(\nabla u_s)
$$

$$
+ \frac{c_\alpha(t - s)}{\zeta_s} \left\{ \Delta \nabla^s u_s \left[ \frac{F^2(\nabla u_s)}{2} \right] - d(\Delta u_s)(\nabla u_s) \right\}
$$

$$
- \frac{\mathcal{N}'(u_s)}{\zeta_s} d\zeta_s(\nabla u_s) - \frac{c_\alpha(t - s)}{2\zeta_s^2} d\zeta_s(\nabla \nabla^s u_s [F^2(\nabla u_s)])
$$

$$
\geq \frac{\mathcal{N}'(u_s)^2}{\zeta_s} F^2(\nabla u_s) + \frac{c_\alpha(t - s)}{\zeta_s F^2(\nabla u_s)} d \left[ \frac{F^2(\nabla u_s)}{2} \right] \left( \nabla \nabla^s u_s \left[ \frac{F^2(\nabla u_s)}{2} \right] \right)
$$

$$
- \frac{\mathcal{N}'(u_s)}{\zeta_s^2} d\zeta_s(\nabla u_s) - \frac{c_\alpha(t - s)}{2\zeta_s^2} d\zeta_s(\nabla \nabla^s u_s [F^2(\nabla u_s)])
$$

in the weak sense. Substituting

$$
d\zeta_s = \frac{1}{\zeta_s} \left\{ \mathcal{N}'(u_s) \mathcal{N}'(u_s) du_s + \frac{c_\alpha(t - s)}{2} d[F^2(\nabla u_s)] \right\}
$$

26
Corollary 4.5

Assume that $(M, F, m)$ is complete and satisfies $\text{Ric}_\infty \geq K > 0$, $C_F < \infty$, $S_F < \infty$ and $m(M) = 1$. Then, for any $u \in C^0_c(M)$ with $0 \leq u \leq 1$ and satisfying (3.6), we have

$$\sqrt{K} \mathcal{N} \left( \int_M u \, dm \right) \leq \int_M \sqrt{K \mathcal{N}^2(u) + F^2(\nabla u)} \, dm. \tag{4.7}$$
Proof. Let \((u_t)_{t \geq 0}\) be the global solution to the heat equation with \(u_0 = u\). Taking \(\alpha = K^{-1}\), we find \(c_n \equiv K^{-1}\) and hence by (4.5)

\[
\sqrt{K} \mathcal{N}^2(u_t) \leq \sqrt{K} \mathcal{N}^2(u_0) + F^2(\nabla u_t) \leq P_{0,t}^{\nabla u} \left( \sqrt{K} \mathcal{N}^2(u) + F^2(\nabla u) \right).
\]

Letting \(t \to \infty\), we deduce from the ergodicity (Proposition 4.3) that

\[
u_t \to \int_M u \, dm,
\]

\[
P_{0,t}^{\nabla u} \left( \sqrt{K} \mathcal{N}^2(u) + F^2(\nabla u) \right) \to \int_M \sqrt{K} \mathcal{N}^2(u) + F^2(\nabla u) \, dm
\]

in \(L^2(M)\). Thereby we obtain (4.7). \(\square\)

4.3 Proof of Theorem 4.1

Proof. Let \(\theta \in (0, 1)\). Fix a closed set \(A \subset M\) with \(m(A) = \theta\) and consider

\[
u^\varepsilon(x) := \max\{1 - \varepsilon^{-1}d(x, A), 0\}, \quad \varepsilon > 0.
\]

Note that \(F(\nabla \nu^\varepsilon) = \varepsilon^{-1}\) on \(B^{-}(A, \varepsilon) \setminus A\), where \(B^{-}(A, \varepsilon) := \{x \in M \mid \inf_{y \in A} d(x, y) < \varepsilon\}\) is the backward \(\varepsilon\)-neighborhood of \(A\). Applying (4.7) to (smooth approximations of) \(\nu^\varepsilon\) and letting \(\varepsilon \downarrow 0\) implies, with the help of \(\mathcal{N}(0) = \mathcal{N}(1) = 0\),

\[
\sqrt{K} \mathcal{N}(\theta) \leq \liminf_{\varepsilon \downarrow 0} \frac{m(B^{-}(A, \varepsilon)) - m(A)}{\varepsilon}.
\]

This is the desired isoperimetric inequality for the reverse Finsler structure \(\overline{F}\) (recall Definition 2.7) since, with \(c := \varphi^{-1}(\theta)/\sqrt{K}\),

\[
\sqrt{K} \mathcal{N}(\theta) = \sqrt{\frac{K}{2\pi}} e^{-Kc^2/2}, \quad \theta = \varphi(\sqrt{K} c) = \sqrt{\frac{K}{2\pi}} \int_{-\infty}^{c} e^{-Kd^2/2} \, da.
\]

Because the curvature bound \(\text{Ric}_\infty \geq K\) is common to \(F\) and \(\overline{F}\), we also obtain (4.1). \(\square\)

References


