CONVEX FUNCTIONALS OF PROBABILITY MEASURES AND NONLINEAR DIFFUSIONS ON MANIFOLDS

KARL-THEODOR STURM

Abstract. The topic of this paper are convexity properties of free energy functionals on the space $\mathcal{P}^2(M)$ of probability measures over a Riemannian manifold. As applications, we obtain contraction properties of nonlinear diffusions on $\mathbb{R}^n$ or on a Riemannian manifold M, regarding them as gradient flows of appropriate free energy functionals. In particular, we present extensions of the Bakry-Emery criterion to nonlinear equations.

1. Introduction and statement of the main results

Throughout this paper, let M be a smooth complete Riemannian manifold of dimension n, with Riemannian distance d and Riemannian volume measure m. We denote by $\mathcal{P}^2(M)$ the space of probability measures on M, equipped with the L^2-Wasserstein distance d_2^W derived from the Riemannian distance on M (see section 3).

Moreover, we fix a lower semicontinuous function $V : M \to \mathbb{R}$ and an increasing function $U : \mathbb{R} \to \mathbb{R}$. We define the *free energy* $S : \mathcal{P}^2(M) \to [-\infty, \infty]$ by

$$S(\nu) := \int_M U\left(\log \frac{d\nu}{dm}\right) d\nu + \int_M V\, d\nu \tag{1.1}$$

provided ν is absolutely continuous w.r.t. the Riemannian volume measure m and $\int U_+(\log \frac{d\nu}{dm})\, d\nu + \int V_+\, d\nu < \infty$ (where we put $U(\log 0) = 0$). Otherwise, we define $S(\nu) := +\infty$.

We say that S is *K-convex* iff

$$\mathrm{Hess}\, S \geq K$$

in some rough sense, to be made precise in section 2. The aim of this paper is to derive conditions on the manifold M as well as on the functions U and V which are necessary and sufficient for K-convextiy of the functional S on $\mathcal{P}^2(M)$.

Remark 1.1. *The importance of K-convexity and our interest in it arises from the fact that K-convexity together with some minimal regularity assumptions on M, U and V imply:*

(i) There exists a unique gradient flow $\sigma : \mathbb{R}_+ \times \mathcal{P}^2(M) \to \mathcal{P}^2(M)$ *for S and it satisfies*

$$d_2^W(\nu_t, \mu_t) \leq e^{-Kt} \cdot d_2^W(\nu_0, \mu_0) \tag{1.2}$$

for all $\nu_0, \mu_0 \in \mathcal{P}^2(M)$ and all $t \geq 0$ where $\nu_t := \sigma(t, \nu_0), \mu_t := \sigma(t, \nu_0)$.

(ii) If in addition $K > 0$ and $\inf S = 0$ then there exists a unique ground state $\nu_\infty \in \mathcal{P}^2(M)$ *satisfying*

$$S(\nu_0) \geq \frac{K}{2} d_2^W(\nu_0, \nu_\infty) \tag{1.3}$$

for all $\nu_0 \in \mathcal{P}^2(M)$. Moreover, along the curves $t \mapsto \nu_t$ of the gradient flow from ν_0 to the ground state ν_∞ we have

$$-\partial_t S(\nu_t) \geq 2K \cdot S(\nu_t) \tag{1.4}$$

1991 *Mathematics Subject Classification.* Primary .

Key words and phrases. gradient flow, nonlinear diffusion, Bakry-Emery criterion, Ricci curvature, entropy, Wasserstein distance, porous medium equation, fast diffusion.

and thus

$$S(\nu_t) \le e^{-2Kt} \cdot S(\nu_0). \tag{1.5}$$

(iii) The curves of the gradient flow are given by $\nu_t(dx) = \rho(t,x)\, m(dx)$ where the densities ρ solve the nonlinear PDE

$$\partial_t \rho(t,x) = \Delta(\rho U'(\log\rho))(t,x) + \nabla(\rho\cdot\nabla V)(t,x) \tag{1.6}$$

on $\mathbb{R}_+ \times M$.

See [OV00], [Vi03], [vRS03], [Ly03], [PePe].

Inequalities (1.3) and (1.4) may be regarded as generalized versions of *Talagrand's inequality* and *Gross' logarithmic Sobolev inequality*, resp. This may be seen in Example 1.2 below where we choose the function S more specifically. If we can verify K-convexity of S for some $K > 0$ then this nonlinear diffusion equation has a unique stationary solution and any other solution converges exponentially fast to the stationary solution.

Example 1.2. The main examples are:

- $U(r) = r, V = 0$ yields the relative entropy $S(\nu) = \int_M \log\frac{d\nu}{dm}\, d\nu$. Its gradient flow is the usual *heat equation* $\partial_t\rho = \Delta\rho$. More precisely, the densities of the gradient flow are solutions of the heat equation.
- $U(r) = r$ leads to the *Fokker-Planck equation*
$$\partial_t\rho = \Delta\rho + \nabla(\rho\cdot\nabla V).$$
In this case, an easy calculation shows $S(\sigma(t,\nu)) = \int u^2\log(u^2)e^{-V}dm$ and $-\partial_t S(\sigma(t,\nu)) = \frac{1}{4}\int|\nabla u|^2 e^{-V}dm$ provided we write $\frac{d\sigma}{dm} = u^2 e^{-V}$. Hence, here we indeed obtain the usual version of the logarithmic Sobolev inequality.
- $U(r) = \frac{1}{a}\exp(ar)$ for a constant $a \neq 0$ and $V = 0$ yields $S(\nu) = \frac{1}{a}\int_M \left(\frac{d\nu}{dm}\right)^a d\nu$. The associated gradient flow is given by the *porous medium equation* (if $a > 0$) or *fast diffusion equation* (if $a < 0$)
$$\partial_t\rho = \Delta(\rho^{1+a}).$$

Our main result yields K-convexity for large classes of energy functionals associated to nonlinear diffusions on Euclidean and Riemannian spaces. As a consequence it yields exponential convergence to equilibrium for the solutions to these equations together with explicit bounds for the rate of convergence.

Theorem 1.3. *Assume that U and V are $\mathcal{C}^2$. Then the free energy S from (1.1) is K-convex if and only if*

$$U''(r) + \frac{1}{n}U'(r) \ge 0 \tag{1.7}$$

and

$$U'(r)\cdot \mathrm{Ric}_x(\xi,\xi) + \mathrm{Hess}_x V(\xi,\xi) \ge K\cdot|\xi|^2 \tag{1.8}$$

for all $r\in\mathbb{R}$, $x \in M$ and $\xi \in T_xM$.

Remark 1.4. The above Theorem has a canonical extension to nonsmooth U and V. Instead of requiring $U \in \mathcal{C}^2$ and (1.7) it suffices to require that the function

$$r \mapsto U(-n\log r) \quad \text{is convex on }]0,\infty[. \tag{1.9}$$

In this case, depending on the sign of Ric the derivative U' in (1.8) should be replaced by the upper or lower derivative, resp.

Instead of restricting to $V \in \mathcal{C}^2$ we may admit any lower semicontinuous function V provided we replace (1.8) by the weaker condition:

$$\underline{\partial}_t^2 V(\gamma_t,\gamma_t) \ge -U'(r)\cdot\mathrm{Ric}(\dot\gamma_t,\dot\gamma_t) + K\cdot|\dot\gamma_t|^2 \tag{1.10}$$

for all $r \in \mathbb{R}$ and all geodesics $\gamma : [0,1] \to M$. For the definition of the lower centered second derivative we refer to section 3.

The *proof* of the above Theorem will be given in sections 4 and 5. Applications of this result to heat equation, Fokker-Planck equation and porous medium equation are straightforward:

Corollary 1.5. *The free energy $S(\nu) = \int_M \log \frac{d\nu}{dm}\, d\nu + \int_M V d\nu$ associated with the Fokker-Planck equation is K-convex if and only if the Bakry-Emery criterion*

$$\mathrm{Ric}_x(\xi,\xi) + \mathrm{Hess}_x V(\xi,\xi) \geq K \cdot |\xi|^2 \tag{1.11}$$

is satisfied ($\forall x \in M, \forall \xi \in T_x M$).

In particular, the relative entropy $S(\nu) = \int_M \log \frac{d\nu}{dm}\, d\nu$ is a K-convex function on the metric space $\mathcal{P}^2(M)$ if and only if the Ricci curvature of the underlying Riemannian manifold M is bounded from below by K.

Corollary 1.6. *For any $N > 0$ the free energy $S(\nu) = -N \cdot \int_M \left(\frac{d\nu}{dm}\right)^{-1/N} d\nu$ associated with the fast diffusion equation $\partial_t \rho = \Delta(\rho^{1-1/N})$ is a convex function on the metric space $\mathcal{P}^2(M)$ if and only if the underlying Riemannian manifold M has nonnegative Ricci curvature and dimension $\leq N$.*

Parts of the above corollaries had been obtained in [OV00], [CMS01] and [vRS03]. The previous results yields a characterization of the curvature-dimension conditions $CD(K,\infty)$ as well as $CD(0,N)$ of Bakry-Emery in terms of contraction properties of nonlinear diffusions. The general condition $CD(K,N)$ may be characterized in a similar manner:

Theorem 1.7. *i) For $K > 0$ and $N > 0$ consider the free energy functional*

$$S(\nu) = \int_M \left[\log\left(\frac{d\nu}{dm}\right) - N \left(\frac{d\nu}{dm}\right)^{-1/N} \right] d\nu \tag{1.12}$$

associated with the nonlinear diffusion equation

$$\partial_t \rho = \Delta(\rho(1 + \rho^{-1/N})).$$

Then S is K-convex if and only if the dimension of the manifold is bounded from above by N and its Ricci curvature is bounded from below by K.

ii) For $K < 0$ and $N > 0$ consider the free energy functional

$$S(\nu) = -N \int_M \log\left[1 + \left(\frac{d\nu}{dm}\right)^{-1/N}\right] d\nu \tag{1.13}$$

associated with the nonlinear diffusion equation

$$\partial_t \rho = \Delta(\rho(1 + \rho^{1/N})^{-1}).$$

Then S is K-convex if and only if the dimension of the manifold is bounded from above by N and its Ricci curvature is bounded from below by K.

Proof. i) The S under consideration corresponds to $U(r) = r - Ne^{-r/N}$. Hence, $U''(r) + \frac{1}{n}U'(r) = \frac{1}{n} + (\frac{1}{n} - \frac{1}{N})e^{-r/N}$ is nonnegative for all $r \in \mathbb{R}$ if and only if $n \leq N$. And $U'(r) \cdot \mathrm{Ric}_x(\xi,\xi) = (1 + e^{-r/N}) \cdot \mathrm{Ric}_x(\xi,\xi$ is bounded from below by $K \cdot |\xi|^2$ (for all r, x, ξ) with some constant $K > 0$ if and only if $\mathrm{Ric}_x(\xi,\xi) \geq K \cdot |\xi|^2$.

ii) In this case, S corresponds to $U(r) = -N \log(1 + e^{-r/N})$. Then $U''(r) + \frac{1}{n}U'(r) = (1 + e^{-r/N})^{-2}(\frac{1}{n} + (\frac{1}{n} - \frac{1}{N})e^{-r/N}$ is nonnegative for all $r \in \mathbb{R}$ if and only if $n \leq N$. And $U'(r) \cdot \mathrm{Ric}_x(\xi,\xi) = (1 + e^{-r/N})^{-1} \cdot \mathrm{Ric}_x(\xi,\xi$ is bounded from below by $K \cdot |\xi|^2$ for all r, x, ξ with some constant $K < 0$ if and only if $\mathrm{Ric}_x(\xi,\xi) \geq K \cdot |\xi|^2$. $\square$

The choice of the functionals S (or the functions U) in the above Theorem is by no means unique. Roughly speaken, for $K>0$ the requirement is that $U(r)\sim r$ as $r\to\infty$ and $U(r)\sim -C\cdot e^{-r/N}$ as $r\to-\infty$. For instance, one could choose

$$U(r)=\begin{cases} r-\beta, & \text{if } r\geq\beta \\ N\left[1-e^{-(r-\beta)/N}\right], & \text{if } r<\beta \end{cases}$$

for any parameter $\beta\in\mathbb{R}$. This leads to $U(\log\rho)=\log_+(\rho/\alpha)+N\left[1-(\rho/\alpha)^{-1/N}\right]_+$ with $\alpha=e^\beta$ and

$$S(\rho\, m)=\int_{\{\rho\geq\alpha\}}\log(\rho/\alpha)\rho\, dm+N\int_{\{\rho<\alpha\}}\left[1-(\rho/\alpha)^{-1/N}\right]\rho\, dm.$$

Similarly, in the case $K<0$ one may choose

$$S(\rho\, m)=\int_{\{\rho\leq\alpha\}}\log(\rho/\alpha)\rho\, dm+N\int_{\{\rho>\alpha\}}\left[1-(\rho/\alpha)^{-1/N}\right]\rho\, dm$$

for any parameter $\alpha>0$. In the latter case, the associated diffusion equation is (at least formally)

$$\partial_t\rho=\Delta(\rho(1\wedge(\rho/\alpha)^{-1/N}))$$

whereas in the former it is

$$\partial_t\rho=\Delta(\rho(1\vee(\rho/\alpha)^{-1/N})).$$

Our results strongly depend on new insights and estimates for the optimal mass transportation on manifolds. From the work of McCann [McC01] and Cordero-Erausquin, McCann, Schmuckenschläger [CMS01] we know that for any pair of absolutely continuous probability measures μ_0 and μ_1 in $\mathcal{P}^2(M)$, there exists a unique geodesic $t\mapsto\mu_t$ in the space $\mathcal{P}^2(M)$ connecting μ_0 and μ_1. Moreover, there exists a vector field Φ such that μ_t is the push forward of μ_0 under the map

$$F_t(x)=\exp_x(t\Phi(x)).$$

It turns out that it is quite important to have precise estimates for the Jacobian dF_t of the map $F_t: M\to M$. The inequality (1.14) below is the key to describe how curvature effects optimal mass transport. It will play a fundamental role in this paper.

Theorem 1.8. *The logarithmic determinant $y_t(x):=\log\det dF_t(x)$ of the Jacobian of F_t satisfies in some appropriate weak sense (to be made precise in Theorem 3.1 below) the following differential inequality in t (for fixed x)*

$$\ddot{y}_t(x)\leq-\frac{1}{n}\dot{y}_t^2(x)-\mathrm{Ric}(\dot{F}_t(x),\dot{F}_t(x)). \tag{1.14}$$

2. K-Convexity

Given an arbitrary geodesic space (N,d_N), a number $K\in\mathbb{R}$ and a function $S:N\to[-\infty,+\infty]$ we say that S is *K-convex* iff for each (constant speed, as usual) geodesic $\gamma:[0,1]\to N$ with $S(\gamma_0)<\infty$ and $S(\gamma_1)<\infty$ and for each $t\in[0,1]$:

$$S(\gamma_t)\leq(1-t)\,S(\gamma_0)+t\,S(\gamma_1)-\frac{K}{2}t(1-t)\,d_N^2(\gamma_0,\gamma_1). \tag{2.1}$$

If S is lower semicontinuous along geodesics, then it suffices to verify this for all geodesics γ and $t=\frac{1}{2}$.

In other words, a function S on a geodesic space N is K-convex if and only if for each geodesic $\gamma:[0,1]\to N$ the function $f:=S(\gamma)$ is K'-convex on $[0,1]$ with $K'=Kd_N^2(\gamma_0,\gamma_1)=K|\dot\gamma|^2$. The function S is called *convex* if it is K-convex for $K=0$. It is called *uniformly convex* if it is K-convex for some $K>0$.

K-convex functions on a interval $I\subset\mathbb{R}$ are semiconvex. Recall that a function $f:I\to\overline{\mathbb{R}}$ is called *semiconvex* iff there exists a smooth function $F:I\to\mathbb{R}$ such that $f+F$ is convex. In particular, semiconvex functions are lower semicontinuous and they are continuous in the interior of the interval $\{f < \infty\} \subset I$. For each semiconvex function $f : I \to \mathbb{R}$ we define the lower centered second derivative

$$\underline{\partial}_t^2 f(t) := \liminf_{s\to 0} \frac{1}{s^2} \cdot [f(t+s) - 2f(t) + f(t-s)]$$

and the centered first derivative

$$\partial_t f(t) := \lim_{s\to 0} \frac{1}{2s} \cdot [f(t+s) - f(t-s)] .$$

The latter limit exists since it may be written as $\frac{1}{2} \cdot \lim_{s\to 0} \frac{1}{s} \cdot [f(t+s) - f(t)] + \frac{1}{2} \cdot \lim_{s\to 0} \frac{1}{s} \cdot [f(t) - f(t-s)]$ and both limits exist, e.g. for convex functions as monotone limits. Analogously, we define semiconcave functions and the upper centered second derivative

$$\overline{\partial}_t^2 f(t) := \limsup_{s\to 0} \frac{1}{s^2} \cdot [f(t+s) - 2f(t) + f(t-s)] .$$

K-convexity is a local property. The above inequality (2.1) holds for a given function S and a given geodesic $\gamma : [0,1] \to N$ provided there exists a partition $0 = t_0 < t_1 < \ldots < t_{n+1} = 1$ such that for each $i = 1, \ldots, n$ the geodesic $\gamma : [t_{i-1}, t_{i+1}] \to N$ satisfies (after suitable reparametrization) inequality (2.1).

A function S is K-convex if and only if it is lower semicontinuous along geodesics and if for each geodesic $\gamma : [0,1] \to N$ with $S(\gamma_0) < \infty$ and $S(\gamma_1) < \infty$ one has $S(\gamma_t) < \infty$ for all $t \in]0,1[$ and

$$\underline{\partial}_t^2 S(\gamma_t) \geq K \cdot d_N^2(\gamma_0, \gamma_1).$$

Example 2.1. A smooth function S on a Riemannian manifold (N, d_N) is K-convex if and only if

$$\operatorname{Hess} S \geq K.$$

Following [McC01], K-convex functions on $N = \mathcal{P}^2(M)$ are also called *displacement K-convex* (to emphasize that it means K-convexity along the geodesics $t \mapsto \gamma_t$ w.r.t. $d_N = d_2^W$ and not along the geodesics $t \mapsto (1-t)\gamma_0 + t\gamma_1$ in the linear space of signed measures).

3. Optimal Mass Transportation on Manifolds

Let us recall some basic results about mass transportation on Riemannian manifolds. Recall that M is a smooth complete Riemannian manifold of dimension n, with Riemannian distance d and Riemannian volume measure m. The set of all probability measures μ on M (equipped with its Borel σ-algebra $\mathcal{B}(M)$) satisfying $\int d^2(x,y)\mu(dy) < \infty$ for some (hence all) $x \in M$ will be denoted by $\mathcal{P}^2(M)$. Given $\mu_0, \mu_1 \in \mathcal{P}^2(M)$ we define their *L^2-Wasserstein distance* by

$$d_2^W(\mu_0, \mu_1) = \inf \left\{ \int_M \int_M d^2(x,y)\pi(dxdy) : \pi \in \mathcal{P}(M^2) \text{ is coupling of } \mu_0 \text{ and } \mu_1 \right\}^{1/2} .$$

Here $\pi \in \mathcal{P}(M^2)$ is called *coupling* (or *transportation plan*) of μ_0 and μ_1 iff its marginals are μ_0 and μ_1, that is, iff $\pi(A \times N) = \mu_0(A)$ and $\pi(M \times A) = \mu_1(A)$ for all $A \in \mathcal{B}(M)$. See e.g. [Du89], [RR98], [Vi03]. The set of absolutely continuous measures in $\mathcal{P}^2(M)$ is a convex subset of $\mathcal{P}^2(M)$ and contains the set $\{S < \infty\}$ for each functional S to be studied in this paper. Hence, in the sequel we may restrict ourselves to absolutely continuous measures in $\mathcal{P}^2(M)$.

Lemma 3.1. *Given two absolutely continuous probability measures $\mu_0 = \rho_0 m$ and $\mu_1 = \rho_0 m$ in $\mathcal{P}^2(M)$ with densities ρ_0, ρ_1 on M, there exists a unique geodesic $t \mapsto \mu_t$ in the space $\mathcal{P}^2(M)$ connecting μ_0 and μ_1. Again each μ_t is absolutely continuous, say $\mu_t = \rho_t m$. Moreover, there exists a vector field Φ such that μ_t is the push forward of μ_0 under the map*

$$F_t(x) = \exp_x(t\Phi).$$

If the measures μ_0, μ_1 are compactly supported then so are all the μ_t for $t \in [0,1]$.

Proof. (i) For absolutely continuous measures with compact support, these results are due to McCann [McC01]. Actually, the vector field is given μ_0-almost everywhere as $\Phi = -\nabla\varphi$ where $\varphi : M \to \mathbb{R}$ shares a some kind of concavity property, called $d^2/2$-concavity (which, however, is entirely different from the notion of concavity used in this paper and introduced in the previous section). Here we only discuss the extension to measures with noncompact support.

(ii) Fix absolutely continuous $\mu_0, \mu_1 \in \mathcal{P}^2(M)$ with some optimal coupling $\pi \in \mathcal{P}(M \times M)$. Given any pair of compact sets $M_0', M_1' \subset M$ with $\pi(M_0' \times M_1') > 0$ define by

$$\pi'(A, B) := \pi(A \cap M_0', B \cap M_1')$$

a coupling $\pi' \in \mathcal{P}(M \times M)$ between the compactly supported, absolutely continuous measures μ_0' and μ_1' defined by

$$\mu_0'(A) := \pi'(A, M), \quad \mu_1'(B) := \pi'(M, B).$$

Both measures μ_0' and μ_1' have total mass $\pi(M_0' \times M_1') \in]0, 1]$. Optimality of π implies optimality of π'. Hence, there exists a vector field Φ' and maps $F_t'(x) = \exp_x(t\Phi')$ such that $\pi' = (id, F_1')_*\mu_0'$ and $t \mapsto \mu_t' := (F_t')_*\mu_0'$ is the unique geodesic connecting μ_0' and μ_1'.

Choosing another pair of compact sets $M_0'' \supset M_0', M_1'' \supset M_1'$ yields a vector field Φ'' such that $F_t''(x) = \exp_x(t\Phi'')$ defines the optimal coupling $\pi'' = (id, F_1'')_*\mu_0''$ and the unique geodesic $t \mapsto \mu_t'' := (F_t'')_*\mu_0''$ between μ_0'' and μ_1''. Uniqueness of the optimal transport now implies that for each t

$$F_t' = F_t'' \quad \mu_0'\text{-a.e. on } M_0'$$

and thus

$$\Phi' = \Phi'' \quad \mu_0'\text{-a.e. on } M_0'.$$

Exhausting $M \times M$ by compact sets $M_0' \times M_1'$ then yields the existence of a vector field Φ and maps $F_t(x) = \exp(t\Phi)$ such that $\pi = (id, F_1)_*\mu_0$ is the unique optimal coupling of μ_0 and μ_1 and $t \mapsto \mu_t := (F_t)_*\mu_0$ is the unique geodesic connecting μ_0 and μ_1. □

Theorem 3.2. *Let $t \mapsto \mu_t = \rho_t m = (F_t)_*\mu_0$ be a geodesic in $\mathcal{P}^2(M)$ connecting two absolutely continuous probability measures μ_0 and μ_1. Then there exists a map $y : M \times [0, 1] \to \mathbb{R}$ such that*

(i) $\forall t \in [0, 1]$: the function $x \mapsto y_t(x)$ is Borel measurable and

$$\rho_0(x) = \rho_t(F_t(x)) \cdot e^{y_t(x)} \qquad \text{for } \mu_0\text{-a.e. } x \in M; \tag{3.1}$$

(ii) $\forall x \in M$: the function $t \mapsto y_t(x)$ is semiconcave (in particular, upper semicontinuous) on $[0, 1]$ and, restricted to $]0, 1[$, it is continuous, has centered derivatives and satisfies:

$$\overline{\partial}_t^2 y_t(x) \le -\frac{1}{n}(\partial_t y_t(x))^2 - \mathrm{Ric}(\dot{F}_t(x), \dot{F}_t(x)). \tag{3.2}$$

Remark 3.3. Equality (3.1) justifies to interpret $J_t(x) := e^{y_t(x)}$ for fixed $t \in [0, 1]$ as the Jacobian determinant $\det dF_t(x)$ of the map $F_t : M \to M$. Indeed, for any measurable function u on M

$$\int_M u\rho_t \, dm = \int_M u(F_t)\rho_0 \, dm = \int_{\{\rho_0 > 0\}} u(F_t)\rho_t(F_t)J_t \, dm.$$

The fundamental inequality (3.2) is closely related to a similar inequality which plays a role in the proof of Bishop-Gromov's volume comparison theorem. Roughly speaken, in the latter result one considers transport problems for measures which are absolutely continuous for the $(n-1)$-dimensional surface measure on spheres and finally obtains an inequality of the form

$$\partial_t^2 y_t(x) \le -\frac{1}{n-1}(\partial_t y_t(x))^2 - \mathrm{Ric}(\dot{F}_t(x), \dot{F}_t(x)), \tag{3.3}$$

cf. [Ch93] (3.42). Note that both, (3.2) and (3.3), are sharp.

We will present two different argumentations for the fundamental inequality (3.2). Firstly, we will give a self-contained and straighforward derivation under the assumption of sufficient smoothness and the absence of cut locus and degeneration effects. Here we will not care about regularity questions (like differentiability of the transport map, existence of conjugate points along the transport rays, nondegeneracy of the Jacobi determinant) but we aim to present the core of the geometric argument. The general case may be deduced from this result using appropriate approximations. This, however, will not be carried out here since we present another proof focussing on the regularity problems. Our second proof works in full generality. It is based on previous calculations and results in [CMS01] and [vRS03].

Proof. Let us fix two absolutely continuous probability measures $\mu_0 = \rho_0 m$ and $\mu_1 = \rho_0 m$ in $\mathcal{P}^2(M)$ with densities ρ_0, ρ_1 on M. Without restriction, we may assume that both are compactly supported. (Otherwise, we have to choose compact exhaustions of $M \times M$ and to consider the restriction of the coupling to these compact sets, see proof of the previous Lemma 3.1). Then there exists a unique geodesic $t \mapsto \mu_t$ in the space $\mathcal{P}^2(M)$ connecting μ_0 and μ_1. Again each μ_t is compactly supported and absolutely continuous, say $\mu_t = \rho_t m$. Moreover, there exists a vector field Φ such that μ_t is the push forward of μ_0 under the map $F_t(x) = \exp_x(t\Phi)$.

First argumentation. Let us for simplicity assume that Φ is a *smooth* vector field and that for each x there are *no conjugate points* on the curve $t \mapsto F_t(x)$, $t \in [0,1]$. For each x, consider the matrix of Jacobi fields

$$\mathcal{A}_t(x) := dF_t(x) : \ T_xM \to T_{F_t(x)}M$$

along the geodesic $F_{\bullet}(x)$. (More precisely, $\mathcal{A}_t(x)v$ is a Jacobi field along $F_{\bullet}(x)$ for each $v \in T_xM$.) It is the unique solution of the Jacobi equation

$$\nabla_t \nabla_t \mathcal{A}_t(x) + R\left(\mathcal{A}_t(x), \dot{F}_t(x)\right)\dot{F}_t(x) = 0 \tag{3.4}$$

with initial conditions $\mathcal{A}_0 = Id$, $\nabla_t \mathcal{A}_0 = \nabla\Phi$. Here R is the curvature tensor and ∇_t denotes covariant derivates along the geodesics $F_{\bullet}(x)$ (cf. [Ch93] (3.4), [Jo95] (4.2.1) or [GHL87] (3.41)). Now assume in addition that the matrix $\mathcal{A}_t(x)$ is *nondegenerate* for all x, t under consideration. In this case, the Jacobi equation immediately implies that the selfadjoint matrix valued map $\mathcal{U} := \nabla_t \mathcal{A} \circ \mathcal{A}^{-1}$ solves the Riccati type equation

$$\nabla_t \mathcal{U}_t + \mathcal{U}_t^2 + R\left(\bullet, \dot{F}_t\right)\dot{F}_t = 0$$

and thus

$$\mathsf{tr}\,(\nabla_t \mathcal{U}_t) + \mathsf{tr}\,(\mathcal{U}_t^2) + \mathrm{Ric}\left(\dot{F}_t, \dot{F}_t\right) = 0. \tag{3.5}$$

Now (cf. [Ch93], Prop. 2.8)

$$\mathsf{tr}\,\mathcal{U}_t = \mathsf{tr}\,(\nabla_t \mathcal{A}_t \circ \mathcal{A}_t^{-1}) = \frac{d}{dt}(\log\det \mathcal{A}_t) = \frac{d}{dt}(\log\det dF_t) = \dot{y}_t.$$

Hence,

$$\mathsf{tr}\,(\mathcal{U}_t^2) \geq \frac{1}{n}(\mathsf{tr}\,\mathcal{U}_t)^2 = \frac{1}{n}(\dot{y}_t)^2 \quad \text{and} \quad \mathsf{tr}\,(\nabla_t \mathcal{U}_t) = \frac{d}{dt}\mathsf{tr}\,(\mathcal{U}_t) = \ddot{y}_t.$$

Together with (3.5), the latter proves inequality (3.2).

Second argumentation. Now we will present an argumentation which holds in the general case, without any restricting smoothness assumptions. Let μ_0, μ_1 be given and define μ_t, ρ_t, F_t as before. Let $\tilde{J}_t := \det dF_t$ be the Jacobian determinant of the map $F_t : M \to M$ (for fixed $t \in [0,1]$) as introduced in [CMS01]. It is well-defined for μ_0-almost all points in M (with exceptional set depending on t). Let M^* be the convex closure of the union of the supports of μ_0 and μ_1 and let K be a lower bound for the Ricci curvature on M^*. Put $\tilde{y}_t(x) = \log \tilde{J}_t(x)$ and

$$\tilde{y}_t^0(x) = \tilde{y}_t(x) - t\tilde{y}_1(x) - \frac{K}{2} \cdot t(1-t) \cdot d^2(x, F_1(x))$$

and define the map $\Theta : \mathcal{B}(M) \times [0,1] \to \mathbb{R}$ by

$$\Theta(A,t) := \int_A \tilde{y}_t^0(x)\,\mu_0(dx).$$

Our first claim is that for each $A \in \mathcal{B}(M)$ the function $t \mapsto \Theta(A,t)$ is concave and satisfies $\Theta(A,0) = \Theta(A,1) = 0$. The latter is obvious by construction. The former was deduced in [vRSt] for the case $A = M$. However, the general case works in the same manner. (In addition, we will give an explicit proof of the more precise estimate (3.2) below.) Representation results for concave functions imply that for each $A \in \mathcal{B}(M)$ there exists a finite measure $\mathbb{P}(A,ds)$ on $[0,1]$ (equipped with its Borel field) such that

$$\Theta(A,t) = \int_0^1 (s \wedge t - st)\,\mathbb{P}(A,ds)$$

for all $t \in [0,1]$. On the other hand, one easily verifies that for each fixed Borel set $B \subset [0,1]$ the map $A \mapsto \mathbb{P}(A,B)$ is a measure, absolutely continuous with respect to μ_0. In other words, $(A,B) \mapsto \mathbb{P}(A,B)$ is a measure on $M \times [0,1]$. Desintegration theory for product measures (or existence of regular conditional expectations) now implies the existence of a kernel $\mathbb{Q} : M \times \mathcal{B}([0,1]) \to \mathbb{R}$ satisfying $\mathbb{P}(A,ds) = \int_A \mathbb{Q}(x,ds)\,\mu_0(dx)$ for all Borel sets $A \subset M$ and thus

$$\Theta(A,t) = \int_0^1 \int_A (s \wedge t - st)\,\mathbb{Q}(x,ds)\,\mu_0(dx).$$

Now put

$$y_t^0(x) := \int_0^1 (s \wedge t - st)\,\mathbb{Q}(x,ds).$$

Then for each $x \in M$ the function $t \mapsto y_t^0(x)$ is concave and for each $t \in [0,1]$ the functions y_t^0 and $\tilde{y}_t^0$ coincide μ_0-a.e. on M. Finally, put

$$y_t(x) := y_t^0(x) + t\tilde{y}_1(x) + \frac{K}{2} \cdot t(1-t) \cdot d^2(x, F_1(x)).$$

Then $t \mapsto y_t(x)$ is semiconcave for each $x \in M$ and for each $t \in [0,1]$ the functions y_t and $\tilde{y}_t = \log \tilde{J}_t$ coincide μ_0-a.e. on M. The change of variable formula for $\tilde{J}_t$ from [CMS] now implies (3.1).

Our next claim is that for each $\epsilon > 0$ there exists a number $s_0 > 0$ such that

$$\frac{1}{s^2}\left[y_{t+s}(x) - 2y_t(x) + y_{t-s}(x)\right] \le \epsilon - \mathrm{Ric}(\dot{F}_t(x), \dot{F}_t(x)) - \frac{n}{s^2}\left(\frac{e^{y_{t+s}(x)/n} - e^{y_{t-s}(x)/n}}{e^{y_{t+s}(x)/n} + e^{y_{t-s}(x)/n}}\right)^2 \tag{3.6}$$

for all $t \in]0,1[$, all $s \in]0,s_0[$ and μ_0-a.e. $x \in M$ (with exceptional set depending on s and t). Choosing an apporiate Borel set M_0 of measure $\mu(M_0) = 0$ we may then redefine the function y on $M_0 \times [0,1]$ (e.g. by $y := 0$) in such a way that the above inequality (3.4) holds for all *rational* $t \in]0,1[$, $s \in]0,s_0]$ and *all* $x \in M$. Since both sides in (3.4) are continuous in s and t it follows that (3.4) holds for *all* s, *all* t and *all* x.

Note that

$$\partial_t y_t(x) = \lim_{s\to 0} \frac{y_{t+s}(x) - y_{t-s}(x)}{2s} = \lim_{s\to 0} \frac{n}{s} \frac{e^{y_{t+s}(x)/n} - e^{y_{t-s}(x)/n}}{e^{y_{t+s}(x)/n} + e^{y_{t-s}(x)/n}}.$$

Hence, in the limit $s \to 0$ inequality (3.4) immediately yields claim (3.2)

$$\overline{\partial}_t^2 y_t(x) \le -\frac{1}{n}(\partial_t y(x))^2 - \mathrm{Ric}(\dot{F}_t(x), \dot{F}_t(x)).$$

In order to prove the inequality (3.4) put $Y_t := e^{y_t(x)/n}$. Then

$$y_{t+s}(x) - 2y_t(x) + y_{t-s}(x) = n\left[\log Y_{t+s} - 2\log(\frac{Y_{t+s}+Y_{t-s}}{2}) + \log Y_{t-s}\right] - 2n\left[\log Y_t - \log(\frac{Y_{t+s}+Y_{t-s}}{2})\right].$$

The first term on the RHS can be estimated from above by

$$n\left(\frac{Y_{t+s}-Y_{t-s}}{Y_{t+s}+Y_{t-s}}\right)^2 = n\left(\frac{e^{y_{t+s}(x)/n}-e^{y_{t-s}(x)/n}}{e^{y_{t+s}(x)/n}+e^{y_{t-s}(x)/n}}\right)^2.$$

In order to estimate the second term, put $d(x) := d(x, F_1(x))$, let $K(x)$ denote a lower bound for the Ricci curvature $\mathrm{Ric}(\dot F_r(x), \dot F_r(x))$ for $r \in [t-s, t+s]$ and let v denote the volume distortion coefficent from [CMS]. Then following [CMS] and [vRSt] we obtain

$$\begin{aligned} -2n\left[\log Y_t - \log(\frac{Y_{t+s}+Y_{t-s}}{2})\right] &\leq -2n\cdot\log v_{\frac12}(F_{t+s},F_{t-s})^{\frac1n} \\ &\leq -2n\cdot\log\left\{\frac{2\sin(\sqrt{\frac{K}{n-1}}sd)}{\sin(\sqrt{\frac{K}{n-1}}2sd)}\right\}^{\frac{n-1}{n}} \\ &\leq s^2\cdot\left[\frac{\epsilon}{2}-K\cdot d^2\right] \\ &\leq s^2\cdot\left[\epsilon-\mathrm{Ric}(\dot F_t(x),\dot F_t(x))\right] \end{aligned}$$

for all $s \leq s_0$ (provided s_0 is sufficiently small). Let us mention that the above estimate only requires to have bounds for the Ricci curvature in direction of the geodesic, and not on the whole space.

Summing up and dividing by s^2 we obtain inequality (3.4). □

Corollary 3.4. *For each $x \in M$, the function $Y_t(x) := e^{y_t(x)/n}$ is semiconcave in $t \in [0,1]$ and satisfies*

$$\overline{\partial}_t^2 Y_t(x) \leq -\frac{1}{n}\mathrm{Ric}(\dot F_t(x), \dot F_t(x))\cdot Y_t(x).$$

This follows immediatley from the above Theorem and the following

Remark 3.5. For any $\mathcal{C}^2$-function u on $\mathbb{R}$

$$\begin{aligned} \frac{1}{s^2}[u(y_{t+s}) &- 2u(y_t)+u(y_{t-s})] \\ &= \frac{1}{s^2}\left[u(y_{t+s})-2u(\frac{y_{t+s}+y_{t-s}}{2})+u(y_{t-s})\right]+\frac{2}{s^2}\left[u(\frac{y_{t+s}+y_{t-s}}{2})-u(y_t)\right] \\ &= u''(\xi_1)\cdot\left(\frac{y_{t+s}-y_{t-s}}{2s}\right)^2+u'(\xi_2)\cdot\frac{1}{s^2}\left[y_{t+s}-2y_t+y_{t-s}\right] \end{aligned}$$

with suitable ξ_1 between y_{t-s} and y_{t+s} and ξ_2 between y_t and $\frac{y_{t+s}+y_{t-s}}{2}$. Similarly, for any convex function u on $\mathbb{R}$

$$\begin{aligned} \frac{1}{s^2}[u(y_{t+s}) &- 2u(y_t)+u(y_{t-s})] \\ &\leq u'_-(\xi)\cdot\frac{1}{s^2}\left[y_{t+s}-2y_t+y_{t-s}\right] \end{aligned}$$

with suitable ξ between y_t and $\frac{y_{t+s}+y_{t-s}}{2}$. Here the left derivative u'_- can also be replaced by the right derivative u'_+.

4. Uniform Convexity of Generalized Entropy Functionals

Let M and m as before, choose an increasing function $U : \mathbb{R} \to \mathbb{R}$ and a lower semicontinuous function $V : M \to \mathbb{R}$ and put

$$S(\rho m) := \int_{\{\rho>0\}} \left[U(\log\rho)+V\right]\rho\, dm$$

for absolutely continuous probability measures $\nu = \rho m$ provided $\int_{\{\rho>0\}} [U_+(\log\rho) + V_+]\,\rho\,dm < \infty$. Otherwise, we define $S(\rho m) := +\infty$. In other words,

$$S(\rho m) = \lim_{k\to\infty} \int_{\{\rho>0\}} \left[U^{(k)}(\log\rho) + V^{(k)}\right] \rho\,dm$$

with $U^{(k)}(r) := U(r) \vee (-k)$, $V^{(k)}(r) := V(r) \vee (-k)$. Hence, without restriction we may assume in the sequel that U and V are bounded from below.

Let us first consider the internal energy

$$S_U(\rho m) := \int_{\{\rho>0\}} U(\log\rho)\rho\,dm$$

Then for each geodesic $t \mapsto \nu_t = \rho_t m$ we obtain

$$\begin{aligned} S_U(\rho_t m) &= \int_{\{\rho_t>0\}} U(\log\rho_t)\,\rho_t\,dm \\ &= \int_{\{\rho_0>0\}} U(\log\rho_t(F_t))\,\rho_t(F_t)\,J_t\,dm \\ &= \int_{\{\rho_0>0\}} U(\log\frac{\rho_0}{J_t})\rho_0 dm \\ &= \int_{\{\rho_0>0\}} U(\log\rho_0 - y_t)\rho_0 dm. \end{aligned}$$

In terms of the functions $U_n(r) := U(-n\log r)$ and $Y_t(x) := \exp(y_t(x)/n)$ this may be rewritten as

$$S_U(\rho_t m) = \int_{\{\rho_0>0\}} U_n(\rho_0^{-1/n}\cdot Y_t)\rho_0 dm.$$

Now let us assume that U_n is convex. Note that for smooth U this is equivalent to (1.7) and for general U it implies that the right and left derivative of U exists and $\inf_r U'(r)$ as well as $\sup_r U'(r)$ are well defined. Moreover, it implies that the map $t \mapsto S_U(\rho_t m)$ is lower semicontinuous (since $t \mapsto Y_t$ is upper semicontinuous and $r \mapsto U_n(r)$ is decreasing and lower semicontinuous).

Our goal is to estimate the second derivative of $t \mapsto S_U(\rho_t m)$. For simplicity we may assume in the following argumentation that U is smooth. Then by Remark 3.5

$$\begin{aligned} \underline{\partial}_t^2 U(\log\rho_0 - y_t) &= \liminf_{s\to0} \frac{1}{s^2}\left[U(\log\rho_0 - y_{t+s}) - 2U(\log\rho_0 - y_t) + U(\log\rho_0 - y_{t-s})\right] \\ &\geq U''(\log\rho_0 - y_t)(\partial_t y_t)^2 - U'(\log\rho_0 - y_t)\overline{\partial}_t^2 y_t \\ &\overset{(1.14)}{\geq} \left\{U''(\log\rho_0 - y_t) + \frac{1}{n}U'(\log\rho_0 - y_t)\right\}(\partial_t y_t)^2 + U'(\log\rho_0 - y_t)\,\mathrm{Ric}(\dot F_t, \dot F_t) \\ &\overset{(1.7)}{\geq} U'(\log\rho_0 - y_t)\,\mathrm{Ric}(\dot F_t, \dot F_t). \end{aligned}$$

Integrating this inequality yields

$$\begin{aligned} &\frac{1}{s^2}\left[U(\log\rho_0 - y_{t+s}) - 2U(\log\rho_0 - y_t) + U(\log\rho_0 - y_{t-s})\right] \\ &\geq \frac{1}{s^2}\int_{-s}^{s} (s - |\xi|)\cdot U'(\log\rho_0 - y_{t+\xi})\,\mathrm{Ric}(\dot{F}_{t+\xi}, \dot{F}_{t+\xi})d\xi \end{aligned}$$

and thus

$$\begin{aligned}
\underline{\partial}_t^2 S_U(\rho_t) &= \liminf_{s\to 0} \frac{1}{s^2}[S_U(\rho_{t+s}m) - 2S_U(\rho_t) + S_U(\rho_{t-s})] \\
&\geq \liminf_{s\to 0} \frac{1}{s^2} \int_M \int_{-s}^{s} (s - |\xi|) \cdot U'(\log\rho_0 - y_{t+\xi}) \operatorname{Ric}(\dot{F}_{t+\xi}, \dot{F}_{t+\xi}) d\xi\, \rho_0\, dm \\
&= \int U'(\log\rho_0 - y_t) \operatorname{Ric}(\dot{F}_t, \dot{F}_t)\, \rho_0\, dm.
\end{aligned}$$

By approximation, this argument extends to all U satisfying (1.9) provided U' is continuous. If one wants to relax the latter, one has to replace U' by bounds for the upper or lower derivative (depending on the sign of the Ricci curvature).

Now let us treat the external energy

$$S_V(\rho m) = \int V \rho dm$$

for lower semicontinuous $V : M \to R$, without restriction assumed to be bounded from below. Then

$$S_V(\rho_t m) = \int V \rho_t dm = \int V(F_t)\rho_t(F_t) J_t dm = \int V(F_t)\rho_0 dm$$

and thus lower semicontinuity of $t \mapsto S_V(\rho_t m)$ is obvious. Moreover,

$$\begin{aligned}
\frac{1}{s^2}[S_V(\rho_{t+s}m) &- 2S_V(\rho_t) + S_V(\rho_{t-s})] \\
&= \frac{1}{s^2} \int [V(F_{s+t}) - 2V(F_t) + V(F_{t-s})]\, \rho_0\, dm.
\end{aligned}$$

This immediately implies

$$\underline{\partial}_t^2 S_V(\rho_t m) \geq \int_M \underline{\operatorname{Hess}}\, V(\dot{F}_t, \dot{F}_t)\rho_0\, dm$$

where for $v \in TM$

$$\underline{\operatorname{Hess}}\, V(v, v) := \liminf_{t\to 0}\ \partial_t^2 V(\exp(tv)).$$

Hence, combining internal and external energy we obtain

$$\underline{\partial}_t^2 S(\rho_t m) \geq \int \Big[U'(\log\rho_0 - y_t) \cdot \operatorname{Ric}(\dot{F}_t, \dot{F}_t) + \underline{\operatorname{Hess}}\, V(\dot{F}_t, \dot{F}_t)\Big]\, \rho_0 dm.$$

Theorem 4.1. *Assume that $V : M \to \mathbb{R}$ is lower semicontinuous, $U : \mathbb{R} \to \mathbb{R}$ is increasing with $U_n : r \mapsto U(-n \log r)$ being convex and that $\forall r > 0$, $\forall v \in TM$:*

$$U'(r) \cdot \operatorname{Ric}(v, v) + \underline{\operatorname{Hess}}\, V(v, v) \geq K \cdot |v|^2.$$

Then $S = S_U + S_V$ is K-convex on $\mathcal{P}^2(M)$: it is lower semicontinuous along geodesics and satisfies

$$\underline{\partial}_t^2 S(\mu_t) \geq K \cdot d^2(\mu_0, \mu_1).$$

Note that in this formulation no second derivative of U is required. Also the existence of the first derivative can be avoided if we interpret U' depending on the sign of Ric as the upper or lower derivative, resp.

Summarizing, we have verified the sufficiency of our conditions for K-convexity in Theorem 1.3.

5. Necessity of the Conditions for K-Convexity

Theorem 5.1. *Let $U : \mathbb{R} \to \mathbb{R}$ be continuous, increasing and let $V : M \to \mathbb{R}$ be continuous. Assume that S as defined in (1.1) is K-convex on $\mathcal{P}^2(M)$ for some $K \in \mathbb{R}$. Then U satisfies (1.9), i.e. $U_n : r \mapsto U(-n \log r)$ is convex.*

Proof. Assume that the function U_n is not convex. Then there exist numbers $r_0, r_1 \in \mathbb{R}$ and $r_{1/2} = \frac{r_0+r_1}{2}$ such that

$$U_n(r_{1/2}) > \frac{1}{2}U_n(r_0) + \frac{1}{2}U_n(r_1).$$

Indeed, this implies that for suitable $\epsilon > 0$

$$U_n(\tilde{r}_{1/2}) > \frac{1}{2}U_n(\tilde{r}_0) + \frac{1}{2}U_n(\tilde{r}_1) + 2\epsilon$$

for all $\tilde{r}_i \in [r_i - \epsilon, r_i + \epsilon]$, $i = 0, 1/2, 1$.

Now fix two points $y, z \in M$, choose $\delta > 0$ sufficiently small (to be specified later), $R \geq \max\{r_0, r_1\}$ and let c_n denote the volume of the unit ball in $\mathbb{R}^n$. Let μ_0 be the Dirac mass in y, let μ_R be the normalized uniform distribution in $B_{\delta R}(y)$ and let $r \mapsto \mu_r$ denote the geodesic in $\mathcal{P}^2(M)$ connecting μ_0 and μ_R. Each of the measures μ_r for $0 < r \leq R$ is absolutely continuous and supported in the ball $B_{\delta r}(y)$. Choosing δ sufficiently small we can achieve that for each $r \in]0, R]$ the density of the measure μ_r is bounded from below by $c_n^{-1}\delta^{-n}(r+\epsilon)^{-n}$ and from above by $c_n^{-1}\delta^{-n}(r-\epsilon)^{-n}$.

Now put $\nu_r = c_n\delta^n\mu_r + (1 - c_n\delta^n)\eta$ where η denotes the normalized uniform distribution in $B_\delta(z)$. Then again $r \mapsto \nu_r$ for $0 \leq r \leq R$ is a geodesic in $\mathcal{P}^2(M)$ and

$$d_W(\nu_{r_0}, \nu_{r_1})^2 = c_n\delta^{n+2}(r_0 - r_1)^2.$$

Moreover, for all $0 < r \leq R$

$$S(\nu_r) \leq c_n \cdot \delta^n \cdot U_n(r - \epsilon) + c_n \cdot \delta^n \cdot \sup_{x \in B_{\delta r}(y)} V(y) + C_\delta$$

and

$$S(\nu_r) \geq c_n \cdot \delta^n \cdot U_n(r + \epsilon) + c_n \cdot \delta^n \cdot \inf_{x \in B_{\delta r}(y)} V(y) + C_\delta$$

where

$$C_\delta = (1 - c_n) \cdot \delta^n \cdot U_n\left(\frac{|B_\delta(z)|^{1/n}}{c_n^{1/n} \cdot \delta}\right) + (1 - c_n \cdot \delta^n) \cdot \frac{1}{|B_\delta(z)|}\int_{B_\delta(z)} V dm$$

(independent of r). Choosing δ sufficiently small we can achieve that

$$\sup_{x \in B_{\delta R}(y)} V(y) \leq \epsilon + \inf_{x \in B_{\delta R}(y)} V(y).$$

Hence, we obtain

$$\begin{aligned} S(\nu_{r_{1/2}}) &- \frac{1}{2}S(\nu_{r_0}) - \frac{1}{2}S(\nu_{r_1}) + \frac{K}{8}d_W^2(\nu_{r_0}, \nu_{r_1}) \\ &\geq c_n\delta^n\left[U_n(r_{1/2} + \epsilon) - \frac{1}{2}U_n(r_0 - \epsilon) - \frac{1}{2}U_n(r_1 - \epsilon) - \epsilon\right] + \frac{K}{8}c_n\delta^{n+2}(r_0 - r_1)^2 \\ &\geq c_n\delta^n[\epsilon - |K|\delta^2(r_0 - r_1)^2/8] > 0 \end{aligned}$$

for δ sufficiently small. This contradicts the K-convexity of S. □

Theorem 5.2. *Let $U \in \mathcal{C}^2(\mathbb{R})$ be increasing and satisfying (1.9) (or equivalently (1.7)) with $n = 1$, let $V \in \mathcal{C}^2(M)$ and $K \in \mathbb{R}$. Assume that S as defined in (1.1) is K-convex on $\mathcal{P}^2(M)$. Then conditon (1.8) is fulfilled.*

Proof. Assume that (1.8) is not true. Then $U'(s)\cdot(\mathrm{Ric}_o(e_1,e_1)+\epsilon)+\mathrm{Hess}_o V(e_1,e_1)+\epsilon\le K$ for some $s\in\mathbb{R}$, some $o\in M$, some unit vector $e_1\in T_oM$ and some $\epsilon>0$. Put $K_1=\mathrm{Ric}_o(e_1,e_1)+\epsilon$ and $K_2=\mathrm{Hess}_o V(e_1,e_1)+\epsilon$.

For $\delta,r>0$ let $A_1:=B_{\delta\exp(-s/n)}(\exp_o(re_1))$ and $A_0:=B_{\delta\exp(-s/n)}(\exp_o(-re_1))$ be geodesic balls of radius $\delta\exp(-s/n)$. Then

$$m(A_0)=c_n\delta^n\exp(-s+O(\delta^2))\quad\text{and}\quad m(A_1)=c_n\delta^n\exp(-s+O(\delta^2))$$

for small δ. As before, c_n denotes the volume of the unit ball in $\mathbb{R}^n$.

Choosing $\delta\ll r\ll 1$ we can find a set $A_{1/2}$ such that $\gamma_{1/2}\in A_{1/2}$ for each minimizing geodesic $\gamma:[0,1]\to M$ with $\gamma_0\in A_0$, $\gamma_1\in A_1$ and such that

$$m(A_{1/2})=c_n\delta^n\exp(-s+O(\delta^2)+r^2/2(K_1-\epsilon/2)+O(r^4))$$

([vRS03], proof of Thm. 1).

Now let μ_0, $\tilde\mu_{1/2}$ and μ_1 be the normalized uniform distribution in A_0, $A_{1/2}$ and A_1, resp. and let $\mu_{1/2}$ be the midpoint in $\mathcal{P}^2(M)$ of μ_0 and μ_1. According to our construction, $\mu_{1/2}$ is supported in the set $A_{1/2}$.

Fix some point z in M with $d(z,o)\gg r$ and put $\nu_i=c_n\delta^n\mu_i+(1-\cdot_n\delta^n)\eta$ for $i=0,1/2,1$ and $\tilde\nu_{1/2}=c_n\delta^n\tilde\mu_{1/2}+(1-c_n\delta^n)\eta$ where as before η denotes the normalized uniform distribution in $B_\delta(z)$. Obviously $\nu_{1/2}$ is the midpoint in $\mathcal{P}^2(M)$ of ν_0 and ν_1. Then for $i=0,1$

$$S_U(\nu_i)=c_n\delta^nU(\log(c_n\delta^n/m(A_i)))+C(\delta)=c_n\delta_n\cdot\left[U(s)+U'(s)\cdot(1+O(\delta^2))\right]+C(\delta)$$

and

$$S_U(\tilde\nu_{1/2})=c_n\delta_n\cdot\left[U(s)+U'(s)\cdot(1+O(\delta^2)+r^2/2(K_1-\epsilon/2)+O(r^4))\right]+C(\delta)$$

where $C(\delta)=(1-c_n)\cdot\delta^n\cdot U\left(\log(c_n\delta^n/m(B_\delta(z)))\right)$.

Now $\nu_{1/2}$ as well as $\tilde\nu_{1/2}$ are supported on the disjoint union of $A_{1/2}$ and $B_\delta(z)$ and they coincide on $B_\delta(z)$. On $A_{1/2}$, the probability measure $\tilde\nu_{1/2}$ has constant density w.r.t. m. Hence.

$$S_U(\nu_{1/2})\ge S_U(\tilde\nu_{1/2}).$$

Indeed, by our assumption on U the function $\psi:t\mapsto\psi(t):=U(\log t)t$ is convex. Hence, if on $A:=A_{1/2}$ the probability measure $\nu_{1/2}$ has density ρ w.r.t. m and if $\tilde\nu_{1/2}$ has constant density α then by Jensen's inequality

$$\int_A U(\log\rho)\rho dm=\int_A\psi(\rho)dm\ge m(A)\cdot\psi\left(\frac{1}{m(A)}\int_A\rho dm\right)=m(A)\cdot\psi(\alpha)=\int_A U(\log\alpha)\alpha\,dm.$$

Hence, for $\delta\ll r\ll 1$

$$\begin{aligned}S_U(\nu_{1/2})-\frac12S_U(\nu_0)-\frac12S_U(\nu_1)&\ge S_U(\tilde\nu_{1/2})-\frac12S_U(\nu_0)-\frac12S_U(\nu_1)\\&\ge c_n\delta^nU'(s)\left[-\frac{K_1}{2}r^2+\frac{\epsilon}{4}r^2+O(r^4)+O(\delta^2)\right]\\&> -U'(s)\frac{K_1}{8}d_2^W(\nu_0,\nu_1)^2.\end{aligned}$$

Now consider S_V.

$$\begin{aligned}S_V(\nu_{1/2})-\frac12S_V(\nu_0)-\frac12S_V(\nu_1)&=c_n\delta^n\int_M\left[V(F_{1/2})-\frac12V(F_0)-\frac12V(F_1)\right]d\mu_0\\&=-c_n\delta^n/2\int_{A_0}\int_0^1[t\wedge(1-t)]\cdot\mathrm{Hess}V(\dot F_t,\dot F_t)\,dt\,d\mu_0\end{aligned}$$

where F_t denotes the transport map pushing forward μ_0 to μ_t. Choosing δ and r small enough one achieves that Hess$V < K_2$ along all transport rays from μ_0 to μ_1. Then

$$S_V(\nu_{1/2}) - \frac{1}{2}S_V(\nu_0) - \frac{1}{2}S_V(\nu_1) > -\frac{K_2}{8}d_2^W(\nu_0,\nu_1)^2.$$

Together with the previous inequality for S_U this yields

$$S(\nu_{1/2}) - \frac{1}{2}S(\nu_0) - \frac{1}{2}S(\nu_1) > -\frac{K}{8}d_2^W(\nu_0,\nu_1)^2$$

which contradicts the K-convexity of S. □

References

[BE84] D. Bakry, M. Emery (1984): Hypercontractivite de semi-groupes de diffusion. *C. R. Acad. Sci., Paris, Ser. I*, **299**, 775–778.

[Ch93] I. Chavel (1993): Riemannian geometry: a modern introduction. Cambridge tracts in mathematics **108**. Cambridge University Press

[CMS01] D. Cordero-Erausquin, R. McCann, M. Schmuckenschläger (2001): A Riemannian interpolation inequality à la Borell, Brascamb and Lieb. *Invent. Math.*, **146**, 219-257.

[Du89] R. M. Dudley (1989): *Real analysis and probability.* Wadsworth & Brooks/Cole.

[GHL87] S. Gallot, D. Hulin, J. Lafontaine (1987): Riemannian geometry. Universitext. Springer-Verlag.

[Jo95] J. Jost (1995): Riemannian geometry and geoemtric analysis. Universitext. Springer-Verlag.

[Ly03] A. Lytchak (2003). Private communication.

[McC01] R. McCann (2001): Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11**, 589-608.

[Ot01] F. Otto (2001): The geometry of dissipative evolution equation: the porous medium equation. *Comm. PDE* **26**, 101-174.

[OV00] F. Otto, C. Villani (2000): Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, **173**, 361–400.

[PePe] Perelman, Petrunin: Unpublished manuscript.

[RR98] S. T. Rachev, L. Rüschendorf (1998): *Mass transportation problems. Vol. I. Theory.* Probability and its Applications. Springer, New York.

[vRS03] M.-K. von Renesse, K.T. Sturm (2003): Transport inequalities, gradient estimates, entropy and Ricci curvature. To appear in Comm. Pure Appl. Math.

[Vi03] C. Villani (2003) *Topics in Mass Transportation.* Graduate Studies in Mathematics. American Mathematical Society.

Institut für Angewandte Mathematik
Universität Bonn
Wegelerstrasse 6
53115 Bonn
Germany
E-mail address: `sturm@uni-bonn.de`