OPTIMAL TRANSPORT AND (RICCI) CURVATURE

Theorems and Problems

(in Search of the geometric meaning of Ricci)

MSRI, 14 Oct 2013

Cédric Villani

University of Lyon
& Institut Henri Poincaré (Paris)

Analytic vs. Synthetic: an example

Def. (i) $\varphi \in C^2(\mathbb{R}^n;\mathbb{R})$ is convex if for any x,

$$\nabla^2\varphi(x) \geq 0$$

Def. (ii) $\varphi : \mathbb{R}^n \to \mathbb{R}$ is convex if for any x, y, t,

$$\varphi\big((1-t)x + ty\big) \leq (1-t)\,\varphi(x) + t\,\varphi(y)$$

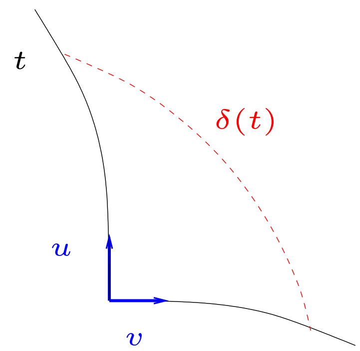
(i): simple, local, effective

(ii): useful, general, stable

– and implies some regularity in the end

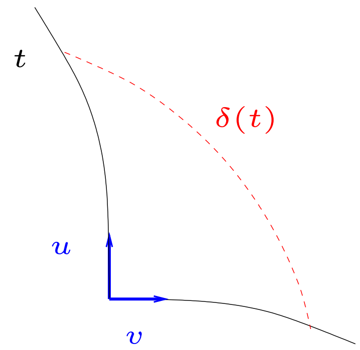
Geometric meaning of curvature

Let $u, v \in T_xM$ be orthogonal unit vectors. $\kappa(u, v)$ measures the divergence of geodesics, w.r.t. to Euclidean geometry: $d(\exp_x tu, \exp_x tv) = \sqrt{2}\,t\,\left(1 - \frac{\kappa}{12}\,t^2 + O(t^4)\right)$



Geometric meaning of curvature

Let $u, v \in T_x M$ be orthogonal unit vectors. $\kappa(u,v)$ measures the divergence of geodesics, w.r.t. to Euclidean geometry: $d(\exp_x tu, \exp_x tv) = \sqrt{2}\, t \left(1 - \frac{\kappa}{12} t^2 + O(t^4)\right)$



Ricci curvature = "average sectional curvature"

$(e, e_2, \ldots, e_n)$ orthonormal, then $\mathrm{Ric}(e) := \sum_{j=2}^{n} \kappa(e, e_j)$

This extends to a quadratic form

(expressed in terms of second derivatives of the metric g)

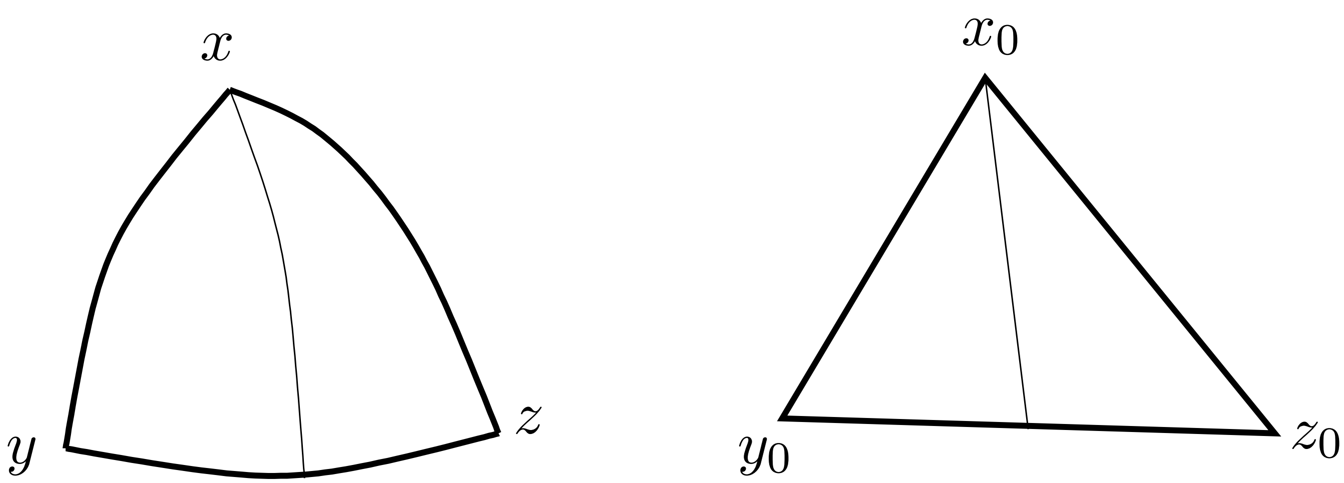
Metric spaces of nonnegative sectional curvature

(Cartan–Alexandrov–Toponogov)

$\kappa \geq 0$

$\Longleftrightarrow$

Triangles are puffier than Euclidean triangles



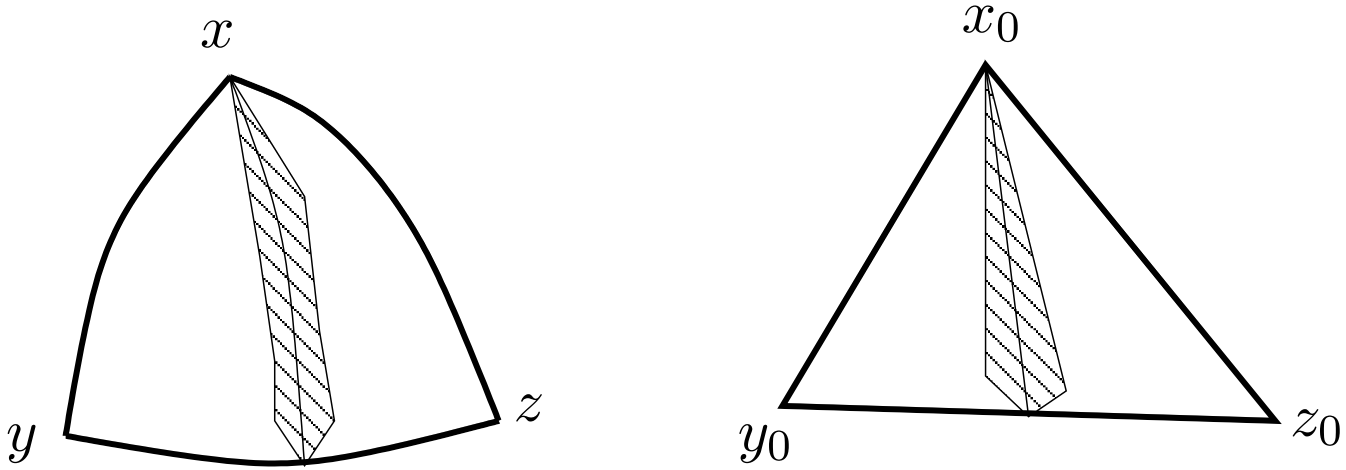
Metric spaces of nonnegative sectional curvature

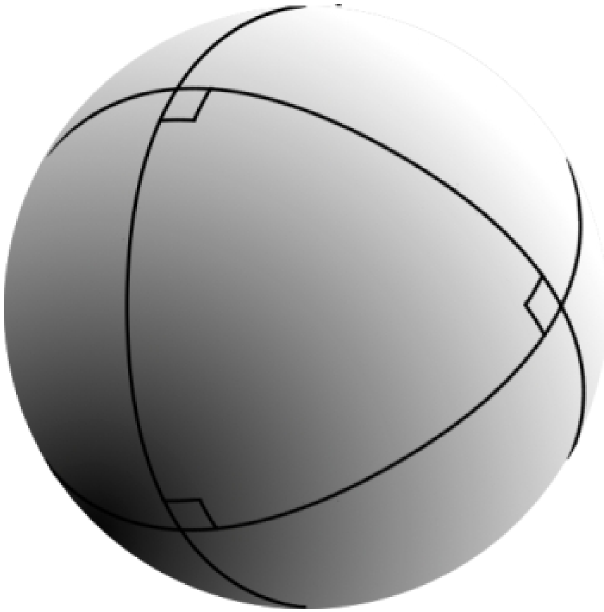
(Cartan–Alexandrov–Toponogov)

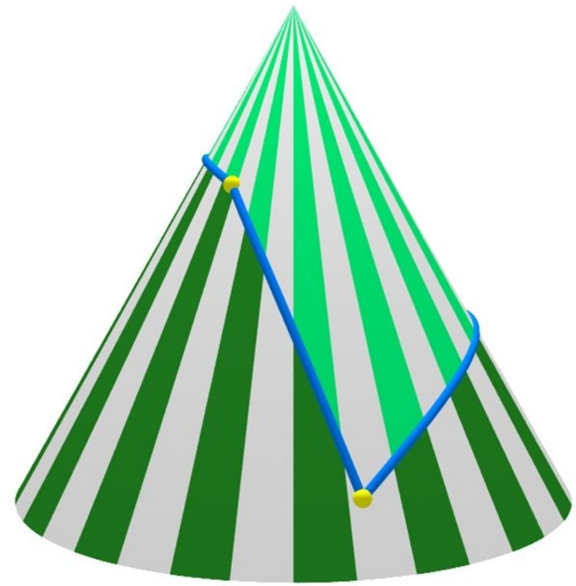
$$\boxed{\kappa \geq 0}$$

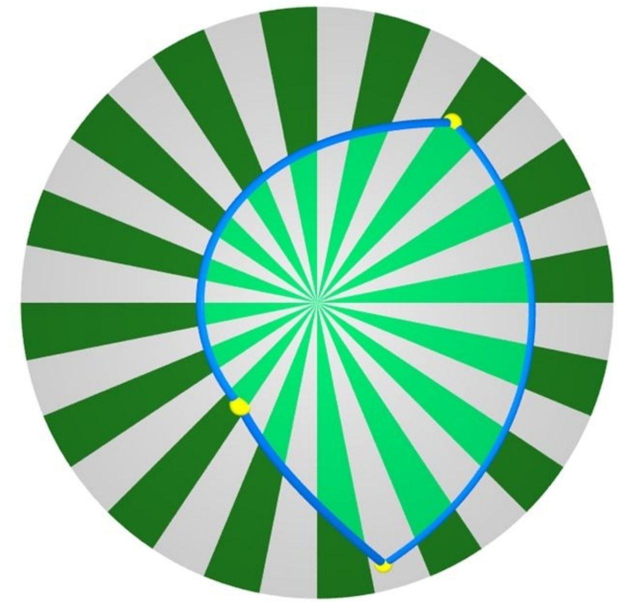
$$\Longleftrightarrow$$

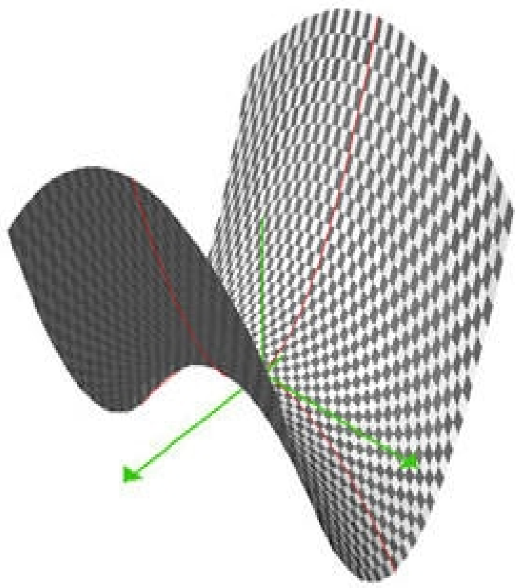
Triangles are puffier than Euclidean triangles

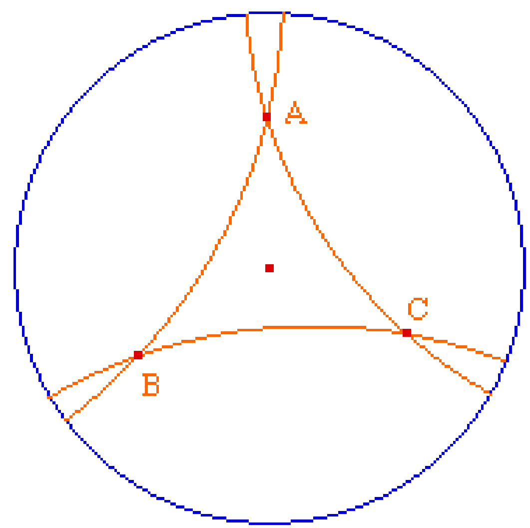


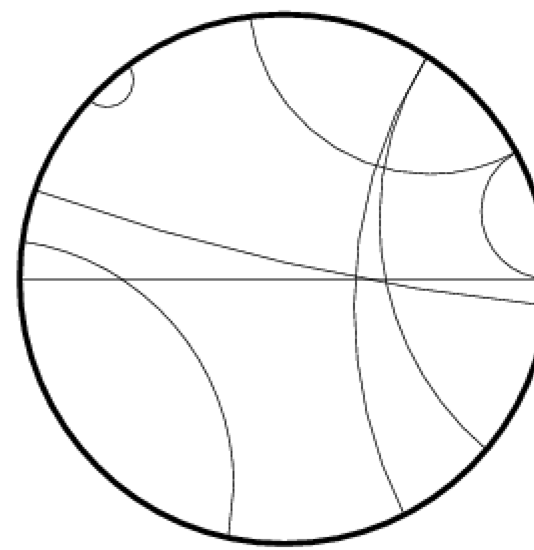









A
B
C



Theory of Alexandrov spaces of positive curvature

Many results by Alexandrov, Burago, Perelman, Petrunin, Ohta, Lytchak, Kuwae, Otsu, Shioya...

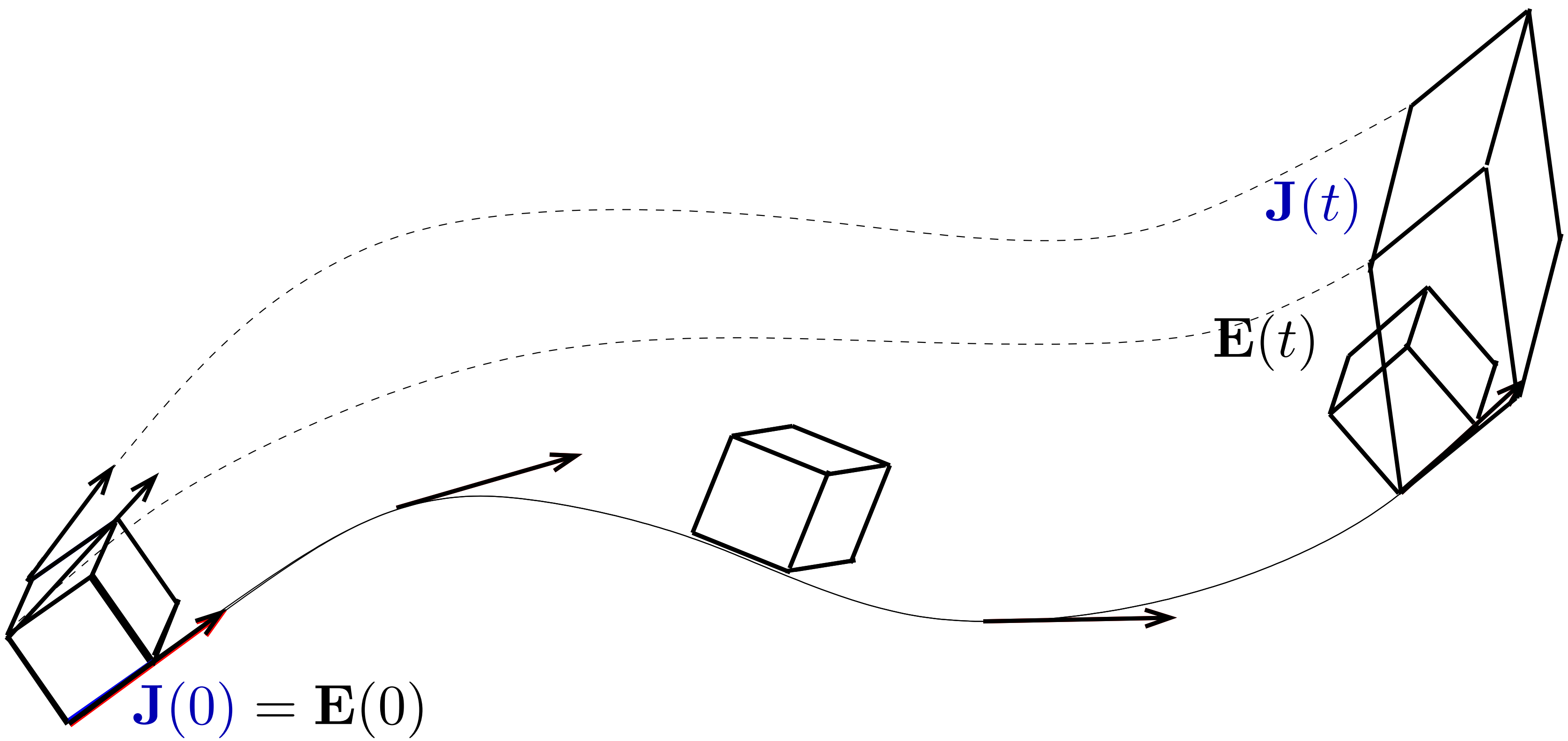
First and second-order calculus, parallel transport, quasigeodesics (replacement for exponential map), basis, gradient flows, smoothing ...

Theory of Alexandrov spaces of negative curvature

Also exists, very different




The meaning of Ricci curvature, I



$$\gamma(t) = \exp_x(tv), \qquad (e_1(t), \ldots, e_n(t)) \qquad e_1(0) = v/|v|$$

$$R_{ij}(t) = \Big\langle \mathrm{Riem}_{\gamma(t)}\big(\dot\gamma(t), e_i(t)\big)\dot\gamma(t), e_j(t) \Big\rangle_{\gamma(t)}$$

$$\ddot{J}(t) + R(t)J(t) = 0$$

$$J(0) = I_n, \qquad \dot{J}(0) = \nabla^2\psi, \qquad \mathrm{tr}\ R = \mathrm{Ric}$$

$$R_{ij}(t) = \left\langle \mathrm{Riem}_{\gamma(\mathrm{t})}\big(\dot\gamma(\mathrm{t}), \mathrm{e_i}(\mathrm{t})\big)\dot\gamma(\mathrm{t}), \mathrm{e_j}(\mathrm{t}) \right\rangle_{\gamma(\mathrm{t})}$$

$$\ddot J(t) + R(t)J(t) = 0$$

$$J(0) = I_n \qquad \dot J(0) = \nabla^2\psi$$

$$\mathcal{J}(t) = \det J(t) \qquad \dot{\mathcal{J}}/\mathcal{J} = \mathrm{tr}\,(\dot J J^{-1}) =: \mathrm{tr}\ U(t)$$

$$\ldots\ \dot U(t) + U(t)^2 + R(t) = 0$$

So $U(t)$ is symmetric!

$$(\mathrm{tr}\ U)^{\cdot} + \mathrm{tr}\ U^2 + \mathrm{Ric} = 0 \implies (\mathrm{tr}\ U)^{\cdot} + \frac{(\mathrm{tr}\ U)^2}{n} + \mathrm{Ric} \le 0$$

$$(\dot{\mathcal{J}}/\mathcal{J})^{\cdot} + n^{-1}(\dot{\mathcal{J}}/\mathcal{J})^2 + \mathrm{Ric} \le 0$$

$$(\mathcal{J}^{\frac{1}{n}})^{\cdot\cdot}(t) \le -\tfrac{1}{n}\,\mathrm{Ric}(\dot\gamma(t))\,\mathcal{J}(t)^{1/n}$$

Lagrangian: If $E(t)$ is an orthonormal matrix of Jacobi fields (infinitesimal geodesic variations of a geodesic γ), then $U := E'E^{-1}$ satisfies the Ricatti equation

$(\mathrm{tr}\, U)' + \mathrm{tr}\,(U^2) + \mathrm{Ric}(\dot\gamma, \dot\gamma) = 0$

Eulerian: If $u \in C^3(M)$, then

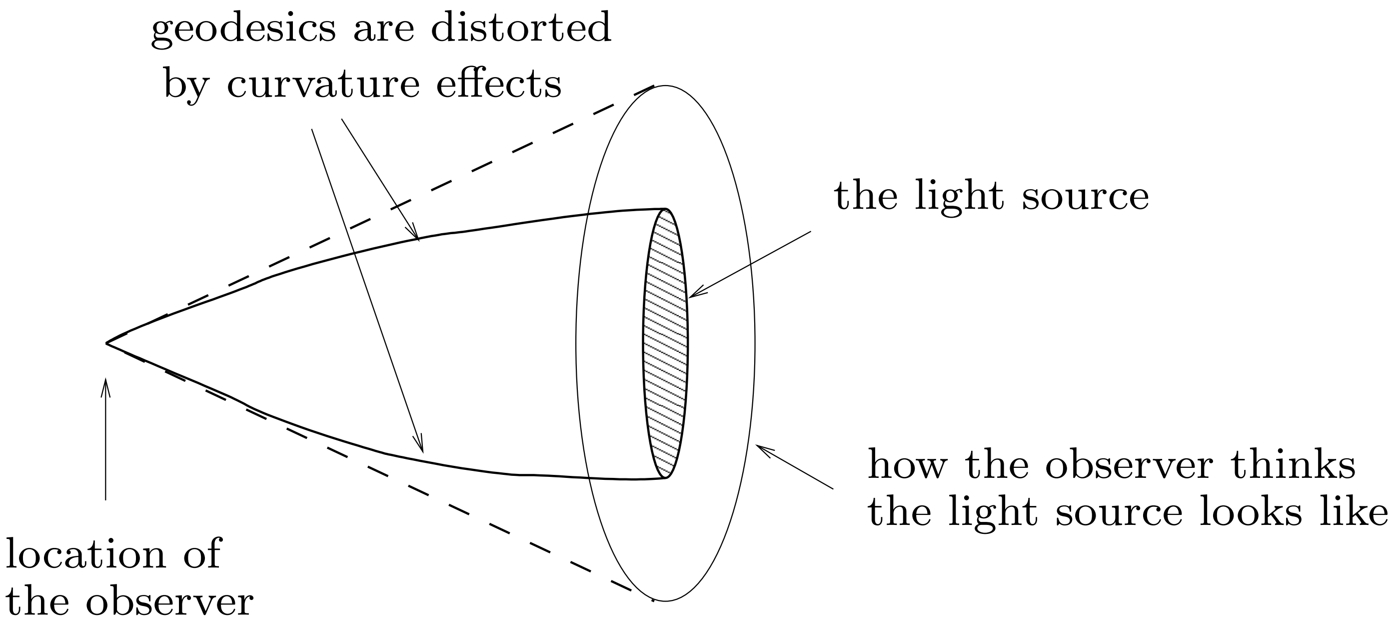
$$-\nabla u \cdot \nabla \Delta u + \Delta \frac{|\nabla u|^2}{2} = \|D^2 u\|^2 + \mathrm{Ric}(\nabla u, \nabla u)$$

Bochner formula

The start of **many** theorems and estimates

Note: Passing from one to the other: Hamilton–Jacobi theory $\partial_t \psi + |\nabla \psi|^2/2 = 0$

The meaning of Ricci curvature, II



Because of nonnegative curvature, the observer overestimates the surface of the light source; in negative curvature this would be the contrary.

$$[\text{Distortion coefficients always} \geq 1] \iff [\mathrm{Ric} \geq 0]$$

The meaning of Ricci curvature, III

Otto's (formal) differential calculus on $P_2(M^n)$, which is the "manifold" of probability measures on M^n, equipped with the distance

$$W_2(\mu_0,\mu_1) = \sqrt{\inf \left\{ \int_0^1 |v(t,x)|^2 \, \mu_t(dx); \qquad \partial_t \mu + \nabla_x \cdot (v\mu) = 0 \right\}}$$

$$= \sqrt{\inf_{T_\# \mu_0 = \mu_1} \int d(x,T(x))^2 \, \mu_0(dx)}$$

The meaning of Ricci curvature, III

Otto's (formal) differential calculus on $P_2(M^n)$, which is the "manifold" of probability measures on M^n, equipped with the distance

$$W_2(\mu_0,\mu_1) = \sqrt{\inf\left\{\int_0^1 |v(t,x)|^2\,\mu_t(dx); \qquad \partial_t\mu + \nabla_x\cdot(v\mu) = 0\right\}}$$

$$= \sqrt{\inf_{T_\#\mu_0=\mu_1}\int d(x,T(x))^2\,\mu_0(dx)}$$

$$\partial_t\rho + \nabla\cdot(\rho\nabla\phi) = 0 \Longrightarrow$$

$$\Big\langle \mathrm{Hess}_{\rho\,\mathrm{vol}}(H)\cdot\partial_t\rho, \partial_t\rho\Big\rangle = \int\Big(\|D^2\phi\|^2 + \langle \mathrm{Ric}\cdot\nabla\phi,\nabla\phi\rangle\Big)\,\rho\,d\mathrm{vol}$$

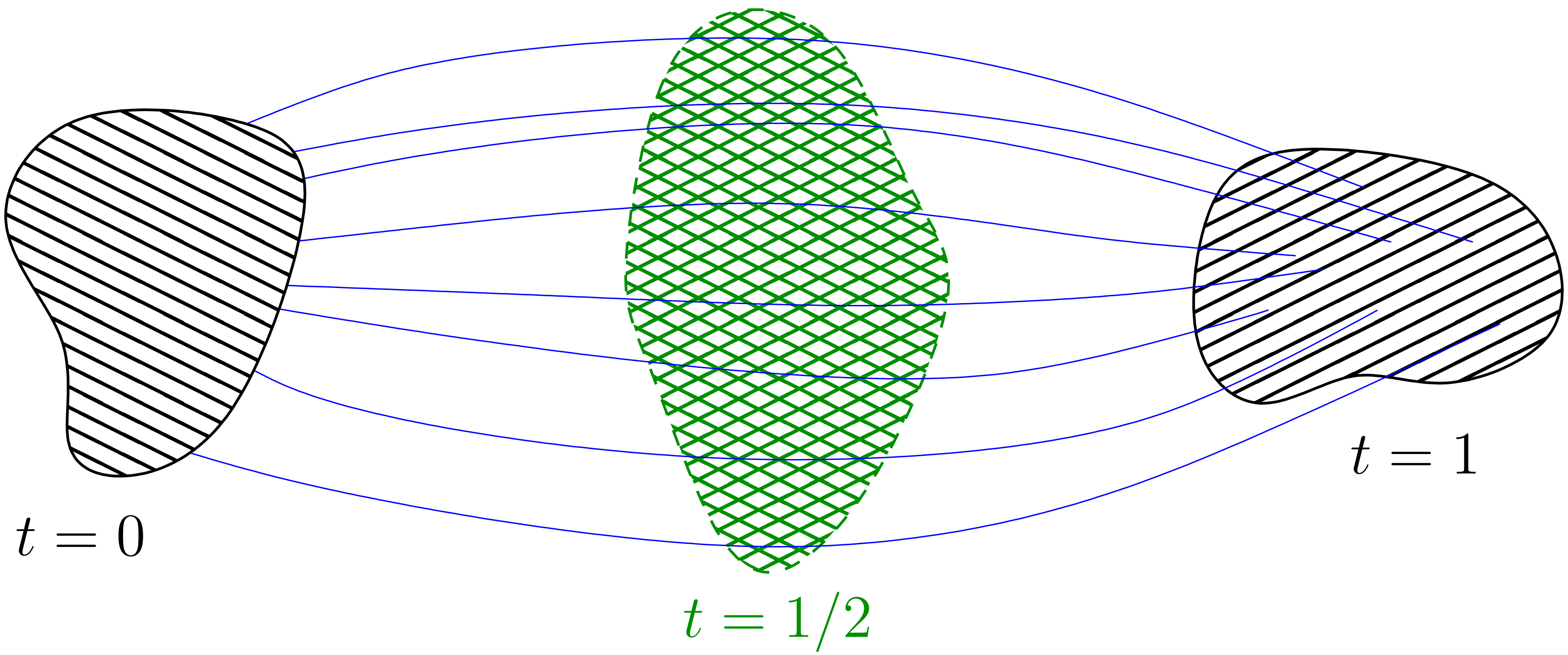
$$H(\rho) = \int \rho\log\rho\,d\mathrm{vol}$$

Consequences of Ricci curvature lower bounds

- isoperimetric inequalities (Lévy–Gromov)
- heat kernel estimates (Li–Yau)
- Sobolev inequalities
- diameter control (Bonnet–Myers)
- spectral gap inequalities (Lichnérowicz)
- Poincaré inequalities (Cheeger...)
- volume growth estimates (Bishop–Gromov)
- compactness of families of manifolds (Gromov)
- concentration (Lévy, Gromov, Talagrand...)
- volume of intermediate points

(Brunn–Minkowski)

Example: curved Brunn–Minkowski for $\mathrm{Ric} \geq 0$



$$\mathrm{vol}^{\frac{1}{n}} m(X,Y) \geq \frac{\mathrm{vol}^{\frac{1}{n}}(X) + \mathrm{vol}^{\frac{1}{n}}(Y)}{2} \qquad \text{(midpoints)}$$

NB: In $\mathbb{R}^n$, recover classical B–M

$|X+Y|^{\frac{1}{n}} \geq |X|^{\frac{1}{n}} + |Y|^{\frac{1}{n}}$ by homogeneity

Inequalities

$$-\nabla u \cdot \nabla \Delta u + \Delta \frac{|\nabla u|^2}{2} = \|D^2 u\|^2 + \mathrm{Ric}(\nabla u)$$

$$\Longrightarrow -\nabla u \cdot \nabla \Delta u + \Delta \frac{|\nabla u|^2}{2} \geq \frac{(\Delta u)^2}{N} + K\,|\nabla u|^2$$

if $n \leq N$, $\mathrm{Ric} \geq K$

$\longrightarrow$ criterion $\mathrm{CD}(K, N)$: **the** inequality involving Ricci curvature and dimension

(Also has Lagrangian counterparts of course)

Generalizations

If the reference volume is $e^{-V(x)}\,\mathrm{vol}(dx)$ then $\mathrm{CD}(K,N)$ ("Ricci $\geq K$, dimension $\leq N$"), should be changed into

$$-\nabla u \cdot \nabla \Delta_\nu u + \Delta_\nu \frac{|\nabla u|^2}{2} \geq \frac{(\Delta_\nu u)^2}{N} + K\,|\nabla u|^2$$

where $\Delta_\nu = \Delta - \nabla V \cdot \nabla$

Equivalently, $\mathrm{Ric}_{N,\nu} \geq K\,g$

where $\mathrm{Ric}_{N,\nu} = \mathrm{Ric} + \nabla^2 V - \dfrac{\nabla V \otimes \nabla V}{N-n}$

Theory of $\mathrm{CD}(K,N)$ bounds

Geometric/analytic consequences have been developed by Bakry, Émery, Ledoux, Li & Yau, and many others.

There it is (often) considered a property of the Laplace operator, or heat equation...

Ex: If $\mathrm{CD}(K,N)$ then

$$\|f\|^2_{L^{\frac{2N}{N-2}}} \leq \|f\|^2_{L^2} + \frac{4}{K\,N\,(N-2)}\|\nabla f\|^2_{L^2}$$

Ex: $\mathrm{CD}(K,\infty)$ corresponds to:

$|\nabla H_t f|^2 \leq e^{-2Kt} H_t |\nabla f|^2$...

Theory of $\mathrm{CD}(K, N)$ bounds

Geometric/analytic consequences have been developed by Bakry, Émery, Ledoux, Li & Yau, and many others.

Gromov: $\mathrm{CD}(K, N)$ can be seen as a property relating distances and volumes, which in the absence of dimension should be considered independent data

Theory of $\mathrm{CD}(K, N)$ bounds

Geometric/analytic consequences have been developed by Bakry, Émery, Ledoux, Li & Yau, and many others.

Gromov: $\mathrm{CD}(K, N)$ can be seen as a property relating distances and volumes, which in the absence of dimension should be considered independent data

Lott–Sturm–Villani

(Building on Otto–Villani, Cordero-Erausquin–McCann-Schmuckenschläger, Sturm–Von Renesse)

The **synthetic** $\mathrm{CD}(K, N)$ criterion goes through the interplay of optimal transport and entropy-type functionals

Recently pushed very far in the "Riemannian" setting

Optimal transport

$$\inf_{T_{\#}\mu_0=\mu_1} \int d(x,T(x))^2\,\mu_0(dx)$$

Boltzmann Entropy

$$S(\rho) = -\int \rho \log \rho = -H(\rho)$$

Optimal transport

<u>**The Kantorovich problem**</u> (Kantorovich, 1942)

- $\mathcal{X}$, $\mathcal{Y}$ two complete separable metric spaces
- $\mu \in P(\mathcal{X})$, $\nu \in P(\mathcal{Y})$
- $c \in C(\mathcal{X} \times \mathcal{Y}; \mathbb{R})$, say $c(x,y) = d(x,y)^2$

$$\Pi(\mu,\nu) = \Big\{\pi \in P(\mathcal{X} \times \mathcal{Y});\ \text{marginals of } \pi \text{ are } \mu \text{ and } \nu\Big\}$$

$$\forall h \quad \int h(x)\,\pi(dx\,dy) = \int h\,d\mu \qquad \int h(y)\,\pi(dx\,dy) = \int h\,d\nu$$

The Kantorovich problem (Kantorovich, 1942)

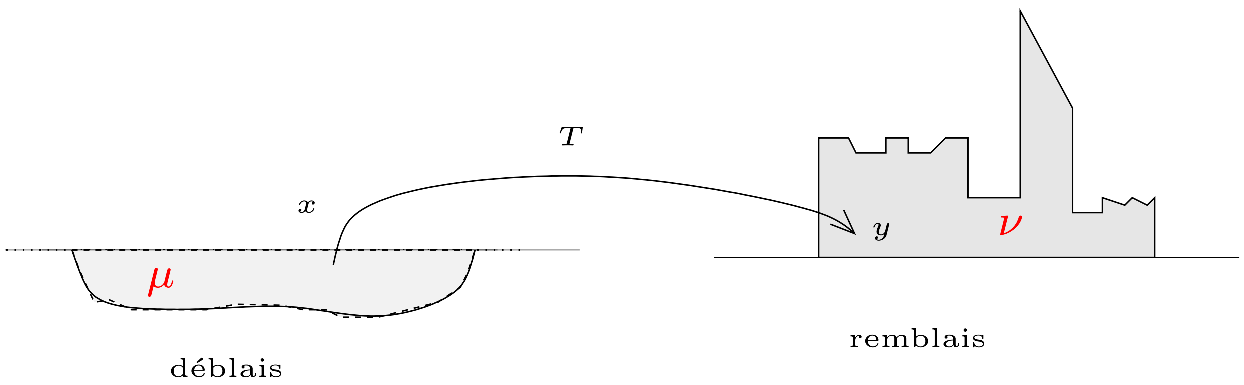
- $\mathcal{X}$, $\mathcal{Y}$ two complete separable metric spaces
- $\mu \in P(\mathcal{X})$, $\nu \in P(\mathcal{Y})$
- $c \in C(\mathcal{X} \times \mathcal{Y}; \mathbb{R})$, say $c(x,y) = d(x,y)^2$

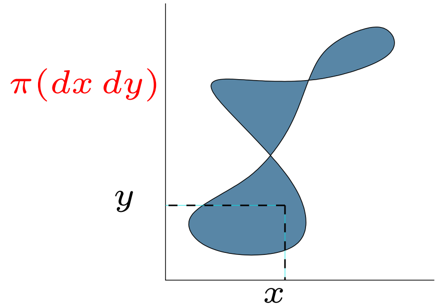
$$\Pi(\mu,\nu) = \Big\{ \pi \in P(\mathcal{X} \times \mathcal{Y});\ \text{marginals of } \pi \text{ are } \mu \text{ and } \nu \Big\}$$

$$\forall h \quad \int h(x)\,\pi(dx\,dy) = \int h\,d\mu \qquad \int h(y)\,\pi(dx\,dy) = \int h\,d\nu$$

$$\textbf{(K)} \quad \boxed{\inf_{\pi \in \Pi(\mu,\nu)} \int c(x,y)\,\pi(dx\,dy)}$$

Engineer's interpretation





Given the initial and final distributions, transport matter at lowest possible cost

Information theory

The Shannon–Boltzmann entropy $S = -\int f \log f$ quantifies how much information there is in a “random” signal Y, or a language.

$$H_\mu(\nu) = \int \rho \log \rho \, d\mu; \qquad \nu = \rho\, \mu.$$

... Entropy = mean value of $\log \dfrac{1}{\rho(Y)}$ ...

Microscopic meaning of the entropy functional

Measures the volume of microstates associated,

to some degree of accuracy in macroscopic observables,

to a given macroscopic configuration (observable distribution function)

$\Longrightarrow$ How exceptional is the observed configuration?

Microscopic meaning of the entropy functional

Measures the volume of microstates associated,

to some degree of accuracy in macroscopic observables,

to a given macroscopic configuration (observable distribution function)

$\Longrightarrow$ How exceptional is the observed configuration?

Boltzmann's formula

$$S = k \log W$$

$\longrightarrow$ How to go from $S = k \log W$ to $S = -\int f \log f$?

Famous computation by Boltzmann

N particles in k boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

N_j = number of particles in box $\#j$

$\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$

$\longrightarrow$ How to go from $S = k \log W$ to $S = -\int f \log f$?

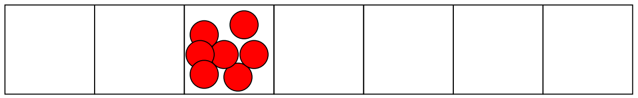
Famous computation by Boltzmann

N particles in k boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

N_j = number of particles in box $\#j$

$\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$



$f = (0, 0, 1, 0, 0, 0, 0)$

$$\boxed{\Omega_N(f) = 1}$$

$\longrightarrow$ How to go from $S = k \log W$ to $S = -\int f \log f$?

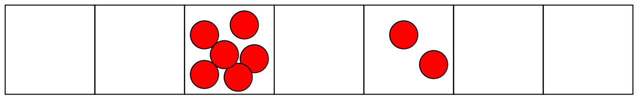
Famous computation by Boltzmann

N particles in k boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

N_j = number of particles in box $\# j$

$\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$



$f = (0, 0, 3/4, 0, 1/4, 0, 0)$

$$\Omega_8(f) = \frac{8!}{6!\, 2!}$$

$\longrightarrow$ How to go from $S = k \log W$ to $S = -\int f \log f$?

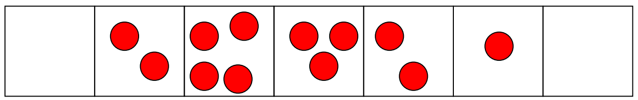
Famous computation by Boltzmann

N particles in k boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

N_j = number of particles in box $\#j$

$\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$



$f = (0, 1/6, 1/3, 1/4, 1/6, 1/12, 0)$

$$\boxed{\Omega_N(f) = \frac{N!}{N_1! \ldots N_k!}}$$

$\longrightarrow$ How to go from $S = k \log W$ to $S = -\int f \log f$?

Famous computation by Boltzmann

N particles in k boxes

$f_1, \ldots, f_k$ some (rational) frequencies; $\sum f_j = 1$

N_j = number of particles in box $\#j$

$\Omega_N(f)$ = number of configurations such that $N_j/N = f_j$

Then as $N \to \infty$

$$\#\Omega_N(f_1, \ldots, f_k) \sim e^{-N \sum f_j \log f_j}$$

$$\frac{1}{N} \log \#\Omega_N(f_1, \ldots, f_k) \simeq -\sum f_j \log f_j$$

Recall: Sanov's Theorem

Mathematical translation of the Boltzmann formula

$x_1, x_2, \ldots$ ("microscopic r.v.") i.i.d. law ν;

$$\widehat{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \text{ (random, "empirical" measure)}$$

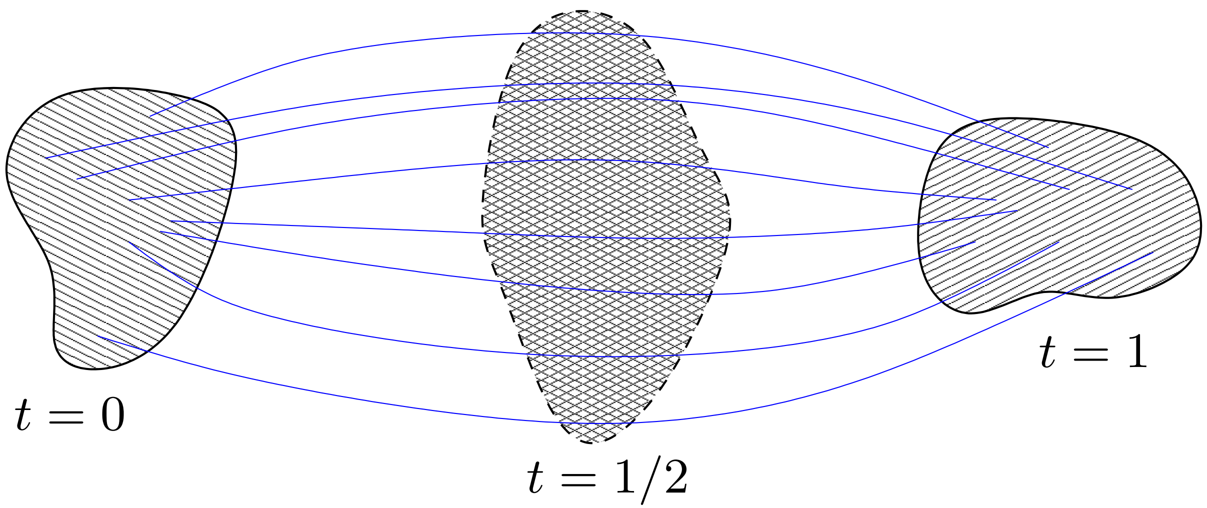
What measure shall we observe??

Informal: $\mathbb{P}\,[\widehat{\mu}^N \simeq \mu] \sim e^{-N H_\nu(\mu)}$

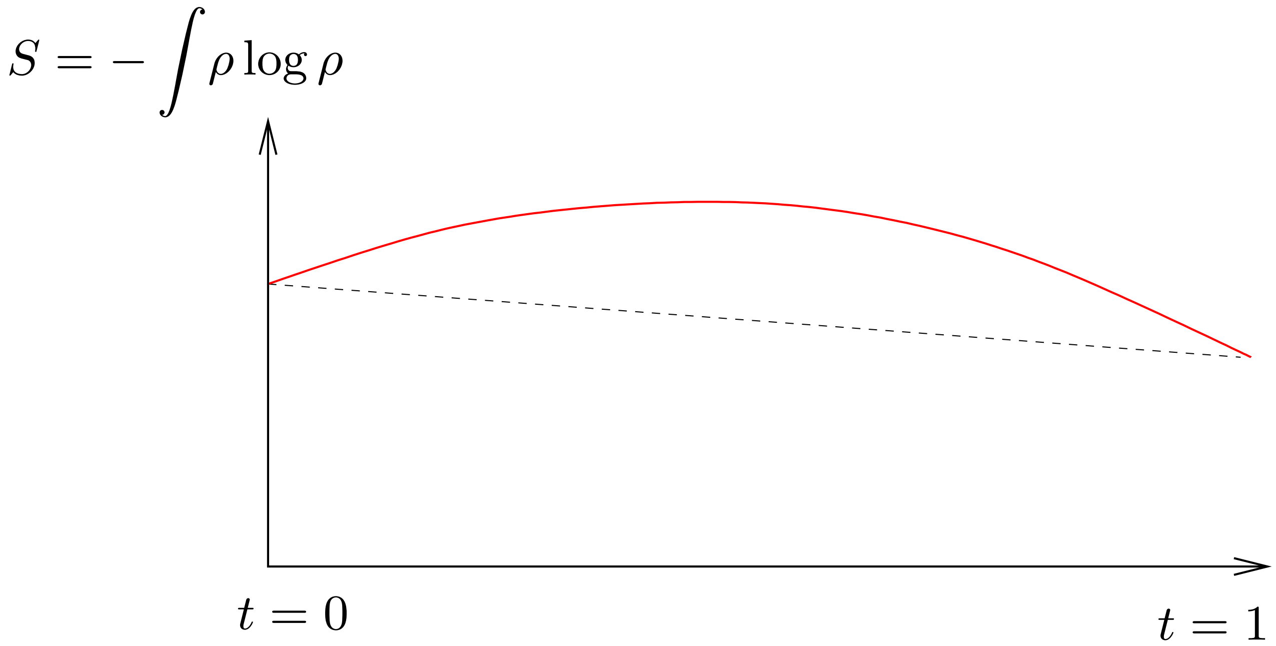
$$H_\nu(\mu) = \int \rho \log \rho \, d\nu, \qquad \rho = \frac{d\mu}{d\nu}$$

Rigorously: H_ν = Large Deviation Rate Function of $\widehat{\mu}^N$.

The lazy gas experiment



$t = 0$
$t = 1/2$
$t = 1$



$S = -\int \rho \log \rho$
$t = 0$
$t = 1$

Relation between transport and Ricci

$$\boxed{\mathrm{Ric} \geq 0}$$

if and only if

$$\boxed{H(\mu_t) = \int \rho_t \log \rho_t \, d\mathrm{vol} \quad \text{is a convex function of } t}$$

$$\rho_t = \frac{d\mu_t}{d\mathrm{vol}}$$

(convexity along geodesics of optimal transport!)


MARC O'POLO
6 7
STOCKHOLM

Metric-measure spaces of nonnegative Ricci curvature

(Lott–Sturm–Villani)

Definition: A compact metric-measured space $(\mathcal{X}, d, \nu)$ has Ricci curvature ≥ 0 (in weak sense) if

$\forall \mu_0, \mu_1 \in P(\mathcal{X}) \quad \exists (\mu_t)_{0 \leq t \leq 1}, \quad$ geodesic in $P(\mathcal{X})$, s.t.

$\forall\, t \in [0,1]$,

$$\int \rho_t \log \rho_t \, d\nu \leq (1-t) \int \rho_0 \log \rho_0 \, d\nu + t \int \rho_1 \log \rho_1 \, d\nu$$

(Some slight variants: a.c. or not? More general $\rho \log \rho$-type nonlinearities?)

Metric-measure spaces of positive Ricci curvature

(Lott–Sturm–Villani)

Definition: A compact metric-measured space $(\mathcal{X}, d, \nu)$ has Ricci curvature $\geq K$ (in weak sense) if

$\forall \mu_0, \mu_1 \in P(\mathcal{X}) \quad \exists (\mu_t)_{0 \leq t \leq 1}, \quad$ geodesic in $P(\mathcal{X})$, s.t.

$\forall\, t \in [0,1]$,

$$\int \rho_t \log \rho_t \, d\nu \leq (1-t) \int \rho_0 \log \rho_0 \, d\nu + t \int \rho_1 \log \rho_1 \, d\nu - \frac{K}{2}\, t(1-t)\, C(\mu, \nu)$$

General $\mathrm{CD}(K,N)$: no unanimity yet!

- Change the class of nonlinearities: in dimension N, replace $\rho \log \rho$ by $U(\rho)$, where $s^N U(s^{-N})$ is convex
- Introduce distortion coefficients in the functional:

$$\int U(\rho_t)\, d\nu \le$$

$$(1-t) \int U\left(\frac{\rho_0(x)}{\beta_t(x,y)}\right) \beta_t(x,y)\, \pi(dy|x)\, \nu(dx) + \ldots$$

where π is optimal, $\beta_t(x,y)$ = reference distortion coeff

Two competing choices of reference distortion coefficients

$$\beta_t(x,y) = \left(\frac{\sin(t\alpha)}{t \sin \alpha}\right)^{N-1}, \qquad \alpha = \sqrt{\frac{K}{N-1}}\, d(x,y) \quad [\mathrm{CD}]$$

$$\beta_t(x,y) = \left(\frac{\sin(t\alpha)}{t \sin \alpha}\right)^{N}, \qquad \alpha = \sqrt{\frac{K}{N}}\, d(x,y) \quad [\mathrm{CD}^*]$$

Consistency

The weak definition coincides with the usual one if the space is smooth (Riemannian manifold)

Core of proof of ($\Rightarrow$) Take $\mu_0 = \rho_0 \,\mathrm{vol}$, $\mu_1 = \rho_1 \,\mathrm{vol}$.

1. The optimal transport takes the following form: each starting point x is related to the final point y by a minimizing geodesic $\gamma_x(t)$, with initial velocity $\dot{\gamma_x}(0) = \nabla\psi(x)$ for some function ψ having some convexity-type properties.

2. The interpolation μ_t between μ_0 and μ_1 is obtained by stopping the geodesic at time t: $\mu_t = (\exp t\nabla\psi)_\#\mu_0$

3. Change variables:

$$H(\mu_t) = H(\mu_0) - \int \log\ \mathrm{Jac}(\exp t\nabla\psi)\, d\mu_0$$

4. $\mathrm{Ric} \geq 0 \Longrightarrow \dfrac{d^2}{dt^2} \log \ \mathrm{Jac}(\exp t\nabla\psi) \leq 0$

.... All the rest is "analysis" and approximation...

Note:

- The entropy is an "integrated" way to involve the logarithmic Jacobian determinant of the exponential map
- With optimal transport we have only access to gradient velocity fields, which is rich enough (Hamilton–Jacobi eq.)

Locality

With the second (weakest) definition of distortion coefficients, the definition is local as soon as the space is nonbranching. Probably the "right" def!

This is because of the underlying differential inequality

$$\ddot{\mathcal{D}} + \frac{K}{N}\mathcal{D} \leq 0$$

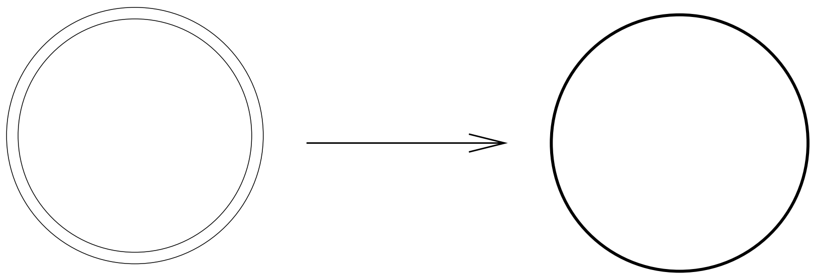
It is open whether the two choices of distortion coefficients are equivalent (true for all examples – cones, Finsler/Alexandrov spaces, warped products...)

It is known that for the nonbranching spaces $\mathrm{CD}^*(K, N)$ is equivalent to Boltzmann's information H satisfying $H'' \geq K + (H')^2/N$ along geodesics

Stability

Def: $(\mathcal{X}_k, d_k, \nu_k)_{k\in\mathbb{N}}$ converges to $(\mathcal{X}, d, \nu)$ in measured Gromov–Hausdorff topology if there are $f_k : \mathcal{X}_k \to \mathcal{X}$

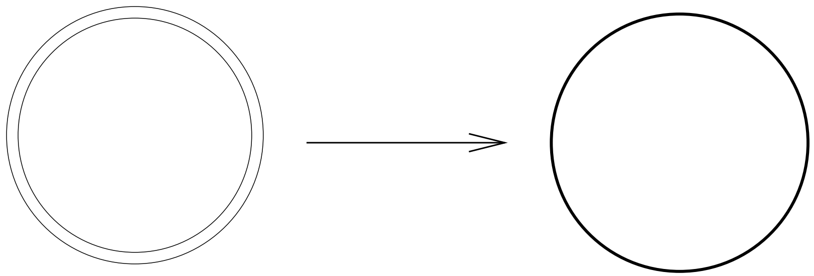
$$\begin{cases} \big| d(f_k(x), f_k(y)) - d_k(x,y) \big| \leq \varepsilon_k \to 0 \\ \forall x \in X, \quad d\big(x, f_k(X_k)\big) \leq \varepsilon_k \\ (f_k)_\# \nu_k \longrightarrow \nu \quad \text{weakly} \end{cases}$$



Stability

Def: $(\mathcal{X}_k, d_k, \nu_k)_{k\in\mathbb{N}}$ converges to $(\mathcal{X}, d, \nu)$ in measured Gromov–Hausdorff topology if there are $f_k : \mathcal{X}_k \to \mathcal{X}$

$$\begin{cases} \big|d(f_k(x), f_k(y)) - d_k(x,y)\big| \leq \varepsilon_k \to 0 \\ \forall x \in X, \quad d\big(x, f_k(X_k)\big) \leq \varepsilon_k \\ (f_k)_\# \nu_k \longrightarrow \nu \quad \text{weakly} \end{cases}$$



Thm: If $(\mathcal{X}_k, d_k, \nu_k)$ has Ric $\geq K$ and converges to $(\mathcal{X}, d, \nu)$ then $(\mathcal{X}, d, \nu)$ has Ric $\geq K$.

(no need for convergence of the second derivatives!)

Strategy of proof of stability (say for $N = \infty$)

Step 1: Reformulate the condition "$\mathrm{Ric} + \nabla^2 V \geq 0$":

For any two probability measures μ_0 *and* μ_1*, there is a geodesic* $(\mu_t)_{0\leq t\leq 1}$ *in the Wasserstein space* $(P(\mathcal{X}), W_2)$*, s.t.* $H_\nu(\mu_t) \leq (1-t)\, H_\nu(\mu_0) + t\, H_\nu(\mu_1)$

Step 2: $P_2(X)$ is stable under MGH:

If $f_k : X_k \to X$ *is an approximate isometry, then* $(f_k)_\# : P_2(X_k) \to P_2(X)$ *also*

Combining with a compactness argument, find a limit geodesic in the space of measures.

Step 3: Use the properties of the entropy to pass to the limit in the inequality.

If $U : \mathbb{R}_+ \to \mathbb{R}_+$ *is convex and continuous, then*

$$U_\nu(\mu) := \int U\left(\frac{d\mu}{d\nu}\right) d\nu$$

is lower semicontinuous w.r.t. μ **and** ν,

and satisfies a contraction principle:

$$\text{for any } f, \qquad U_{f_\# \nu}(f_\# \mu) \leq U_\nu(\mu)$$

Conclude that the same property holds true in the limit space, deduce $\mathrm{Ric} + \nabla^2 V \geq 0$.

Compatibility (Petrunin 2009)

If $(\mathcal{X}, d)$ is a compact finite-dimensional Alexandrov space with "sectional" curvature ≥ 0 then also $(\mathcal{X}, d, \mathrm{vol})$ has "Ricci" curvature ≥ 0.

This establishes a direct link between Cartan–Alexandrov–Toponogov and Lott–Sturm–V and ensures the compatibility of weak definitions

This was generalized to "sectional curvature $\geq \kappa$", providing examples of $\mathrm{CD}(K, N)$ spaces.

But weak $\mathrm{CD}(K, N)$ spaces are more general and include all MGH limits of $\mathrm{CD}(K, N)$ manifolds, all normed $\mathbb{R}^N$...

Properties derived from the synthetic formulation

Sobolev inequalities, Brunn–Minkowski, Bishop–Gromov, Poincaré, Lichnérowicz...

Example: Prove the Curved Brunn–Minkowski inequality

A_0, A_1 given

$$\mu_0 := \nu|_{A_0}, \quad \mu_1 := \nu|_{A_1}; \qquad (\mu_t)_{0 \le t \le 1}$$

$$\int \rho_{1/2}^{1-1/N} \, d\nu \ge \frac{1}{2}\left(|A_0|^{1/N} + |A_1|^{1/N}\right),$$

Apply Jensen to conclude.

Isoperimetric inequalities, concentration

The transport approach gives a grip on measures/sets

Used for concentration inequalities

Recently used by Funano to prove: under $\mathrm{CD}(0,\infty)$, $\lambda_k(M,\nu) \leq C^k \lambda_1(M,\nu)$ for some universal C.

The key is the entropy interpretation and a recursive estimate on the **separation**: $\mathrm{Sep}(M,\nu,\alpha_1,\ldots,\alpha_N) :=$ maximum min-distance between sets $A_1,\ldots,A_N$ satisfying $\nu[A_j] = \alpha_j$,

obtained through displacement convexity of H

Also inequalities on isoperimetric-type constants...

Isoperimetric inequalities, concentration

The transport approach gives a grip on measures/sets

Used for concentration inequalities

Recently used by Funano to prove: under $\mathrm{CD}(0,\infty)$, $\lambda_k(M,\nu) \leq C^k \lambda_1(M,\nu)$ for some universal C.

The key is the entropy interpretation and a recursive estimate on the **separation**: $\mathrm{Sep}(M,\nu,\alpha_1,\ldots,\alpha_N) :=$ maximum min-distance between sets $A_1,\ldots,A_N$ satisfying $\nu[A_j]=\alpha_j$,

obtained through displacement convexity of H

Also inequalities on isoperimetric-type constants...

Why does the Lévy–Gromov inequality remain elusive?

Rough heat flow (Ambrosio–Gigli–Savaré 2011)

If $(\mathcal{X}, d, \nu)$ has "Ricci" curvature $\geq -K$, one can define a (nonlinear) heat flow on the space of probability densities,

- either as gradient flow of H_ν in P_2
- or as L^2 grad flow of Dirichlet form $\int |\nabla \rho|^2 \, d\nu$

Origin: Jordan–Kinderlehrer–Otto (1998)

On M compact Riemannian manifold (or $M = \mathbb{R}^n$) there is a link between

- heat/Fourier equation $\dfrac{\partial \rho}{\partial t} = \Delta \rho$ on M
- Boltzmann's H functional: $H(\rho) = \displaystyle\int \rho \log \rho$
- optimal transport

$$C(\mu, \nu) = \inf_{T_{\#}\mu = \nu} \int d\big(x, T(x)\big)^2 \, \mu(dx)$$

Monge solution of Fourier equation

Unorthodox gradient flow scheme. Time discretize.

From time t to time $t + \Delta t$: Given $\rho(t)$, search for

$\rho(t + \Delta t)$ as the minimizer of $H(\rho) + \dfrac{C(\rho(t), \rho)}{2\,\Delta t}$

Monge solution of Fourier equation

Unorthodox gradient flow scheme. Time discretize.

From time t to time $t + \Delta t$: Given $\rho(t)$, search for $\rho(t + \Delta t)$ as the minimizer of $H(\rho) + \dfrac{C(\rho(t), \rho)}{2\,\Delta t}$

Note: Interpretation (Peletier et al.)

The occurrence of H can be related to Sanov's Theorem; the exponent 2 to the (log) Central Limit Theorem

Theorem (Ambrosio–Gigli–Savaré)

This procedure works the same in weak $\mathrm{CD}(K, \infty)$ spaces. In fact as soon as $|\nabla^- H_\nu|$ is lower semi-continuous; in that case this is the square root of the Fisher information.

Side PDE remark

Nonsmooth Hamilton–Jacobi theory is crucial here!

For this purpose it was developed in general metric spaces (Lott, V, Gozlan, Roberto, Samson, Ambrosio, Gigli, Savaré...)

$$Q_t f(x) = \min \left[f(y) + \tfrac{d(x,y)^2}{2t} \right]$$

$\Longrightarrow$ In any geodesic space, $\partial_t Q_t f + \dfrac{|\nabla Q_t f|^2}{2} = 0$ (except at countably many times)

How wide is this generalization?

Nonbranching $\mathrm{CD}(K, N)$ spaces satisfy many properties of smooth ones.

But the flow is in general nonlinear and the splitting theorem does not hold; normed spaces are allowed, Finsler geometry is included

$\mathrm{RCD}(K,N)$ **Spaces** / $\mathrm{RCD}^*(K,N)$ **Spaces**

If one makes the **additional assumption** that $W^{1,2}$ is Hilbert (Ambrosio–Gigli–Savaré), or equivalently (!) that the heat flow is linear, then one obtains a narrower class of weak $\mathrm{CD}(K,N)$ spaces, which satisfies a lot, and is still stable (!)

- Laplace operator
- Bochner inequality; link to Bakry–Émery formalism (equivalence: Erbar–Kuwada–Sturm); cone property, etc.
- Splitting Theorem (Gigli), in quantitative form; sharp inequalities for RCD^*
- Almost everywhere existence of (unique?) finite-dimensional tangent spaces

An intermediate theory?

Can one develop a good calculus without restricting to the "Riemannian" assumption, keeping Finsler spaces along the way?

Maybe if the Sobolev space $W^{1,2}$ is strictly convex?

Note: Exponent 2 is there in the heat equation (even nonlinear) and in the curvature!

Adaptation to discrete spaces

Many different theories in discrete spaces (approximate geodesics, or change the distance through a discretized Riemannian structure, etc.)

Ollivier, Sturm–Bonciocat, Maas, Erbar, Mielke, Gozlan–Roberto–Samson–Tetali, Hillion...

The Ricci curvature of the discrete hypercube?

(Question by D. Stroock, 1998)

Ollivier–V: A and B two nonempty subsets of $\{0,1\}^N$. M the set of midpoints of A and B. Then

$$\log |M| \geq \frac{1}{2}\big(\log |A| + \log |B|\big) + \frac{K}{8} d(A,B)^2, \qquad K = \frac{1}{2N}$$

Maas: Can be made more precise, $K = 1/(2N)$ is (in some sense) the discrete Ricci curvature of the hypercube.