SYNTHETIC THEORY
OF RICCI CURVATURE

WHEN INFORMATION THEORY, OPTIMIZATION,
GEOMETRY AND GRADIENT FLOWS MEET

COLT, 6 July 2015

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) $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if for any $x, y, t$,

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

(i): simple, local, effective

(ii): useful, general, stable

– and implies some regularity in the end
Geometric meaning of curvature

Let $u, v \in T_x M$ be orthogonal unit vectors. $\kappa(u, v)$ measures the divergence of geodesics, w.r.t. to Euclidean geometry:

$$d(\exp_x tu, \exp_x tv) = \sqrt{2} t \left( 1 - \frac{\kappa}{12} t^2 + O(t^3) \right)$$
**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:

$$d(\exp_x tu, \exp_x tv) = \sqrt{2} t \left( 1 - \frac{\kappa}{12} t^2 + O(t^3) \right)$$

**Ricci curvature = “average sectional curvature”**

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

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)

$$\kappa \geq 0$$

$\iff$

Triangles are 

**puffer**

than Euclidean triangles
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
\begin{align*}
\gamma(t) &= \exp_x(tv), \\
R_{ij}(t) &= \left< \text{Riem}_\gamma(t)(\dot{\gamma}(t), e_i(t))\dot{\gamma}(t), e_j(t) \right> \\
\dot{\gamma}(t), e_i(t), e_j(t) &
\end{align*}
\( R_{ij}(t) = \left\langle \text{Riem}_{\gamma(t)}(\dot{\gamma}(t), e_i(t)) \dot{\gamma}(t), e_j(t) \right\rangle_{\gamma(t)} \)

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

\( J(0) = I_n \quad \dot{J}(0) = \nabla^2 \psi \)

\( \mathcal{J}(t) = \det J(t) \quad \dot{\mathcal{J}} / \mathcal{J} = \text{tr} (\dot{J} J^{-1}) =: \text{tr} U(t) \)

\[ \dot{U}(t) + U(t)^2 + R(t) = 0 \quad \text{So } U(t) \text{ is symmetric!} \]

\( (\text{tr } U) \cdot + \text{tr } U^2 + \text{Ric} = 0 \quad \Rightarrow \quad (\text{tr } U) \cdot + \frac{(\text{tr } U)^2}{n} + \text{Ric} \leq 0 \)

\[ (\dot{\mathcal{J}} / \mathcal{J}) \cdot + n^{-1}(\dot{\mathcal{J}} / \mathcal{J})^2 + \text{Ric} \leq 0 \]

\( (\mathcal{J}^\frac{1}{n}) \cdot (t) \leq -\frac{1}{n} \text{Ric}(\dot{\gamma}(t)) \mathcal{J}(t)^{1/n} \)
Lagrangian: If \( E(t) \) is an orthonormal matrix of Jacobi fields (infinitesimal geodesic variations of a geodesic \( \gamma \)), then \( U := E'E^{-1} \) satisfies the Ricatti equation
\[
(tr\ U)' + tr\ (U^2) + Ric(\dot{\gamma}, \dot{\gamma}) = 0
\]

Eulerian: If \( u \in C^3(M) \), then Bochner formula:
\[
-\nabla u \cdot \nabla \Delta u + \Delta \frac{||\nabla u||^2}{2} = ||D^2 u||^2 + Ric(\nabla u, \nabla u)
\]
\[
\implies -\nabla u \cdot \nabla \Delta u + \Delta \frac{||\nabla u||^2}{2} \geq \frac{(\Delta u)^2}{N} + K||\nabla u||^2
\]
CD\((K, N)\) estimate ("Ric \geq K, dim \leq N")
The start of many theorems and estimates!

Note: Passing from one to the other: Hamilton–Jacobi theory \( \partial_t \psi + |\nabla \psi|^2/2 = 0 \)
Because of nonnegative curvature, the observer overestimates the surface of the light source; in negative curvature this would be the contrary.

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

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

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

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

$$= \sqrt{\inf \int d(x, T(x))^2 \mu_0(dx)}$$

$$\partial_t \rho + \nabla \cdot (\rho \nabla \phi) = 0 \implies \left\langle \Hess_{\rho \text{vol}}(H) \cdot \partial_t \rho, \partial_t \rho \right\rangle = \int \left( \|D^2 \phi\|^2 + \langle \Ric \cdot \nabla \phi, \nabla \phi \rangle \right) \rho \, d\text{vol}$$

$$H(\rho) = \int \rho \log \rho \, d\text{vol}$$
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 $\text{Ric} \geq 0$

$$\text{vol}^{\frac{1}{n}} m(X, Y) \geq \frac{\text{vol}^{\frac{1}{n}}(X) + \text{vol}^{\frac{1}{n}}(Y)}{2}$$  \quad \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
Generalizations

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

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

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

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

where $\text{Ric}_{N,\nu} = \text{Ric} + \nabla^2 V - \frac{\nabla V \otimes \nabla V}{N - n}$
Theory of 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 CD\((K, N)\) then

\[
\|f\|_{L^2}^{2N} \leq \|f\|_{L^2}^2 + \frac{4}{KN(N-2)}\|\nabla f\|_{L^2}^2
\]

Ex: CD\((K, \infty)\) corresponds to:

\[
|\nabla H_t f|^2 \leq e^{-2Kt}H_t |\nabla f|^2...
\]
Theory of $\text{CD}(K, N)$ bounds

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

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

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

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

**Synthetic theory??**
Gradient flows
\[ \dot{X} = -\nabla \mathcal{E}(X) \]

Boltzmann Entropy
\[ S(\rho) = -\int \rho \log \rho = -H(\rho) \]

Optimal transport
\[ \inf_{T_\#\mu_0=\mu_1} \int d(x, T(x))^2 \mu_0(dx) \]
Gradient flows in a nonsmooth setting

For instance, time-discretize

\[ X^n_\tau \rightarrow X^\tau_{n+1} \]

\[ X_{n+1} \text{ sol of min } \left[ \mathcal{E}(X) + \frac{d(X_n, X)^2}{2\tau} \right] \]

\[ n \rightarrow \infty, \tau \rightarrow 0, n\tau \rightarrow t, \]

\[ X^n_\tau \simeq X(t) \]
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.

\[
H_{\nu}(\mu) = \int \rho \log \rho \, d\nu; \quad \mu = \rho \nu.
\]

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

⇒ 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)

⇒ How exceptional is the observed configuration?

Boltzmann’s formula

\[ S = k \log W \]
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 $N_j/N = f_j$
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; $\sum f_j = 1$

$N_j = \text{number of particles in box } \#j$

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

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

$$\Omega_N(f) = 1$$
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; $\sum f_j = 1$

$N_j = \text{number of particles in box } \#j$

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

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

\[
f = (0, 0, 3/4, 0, 1/4, 0, 0)
\]
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 $N_j/N = f_j$

\[
\begin{pmatrix}
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot
\end{pmatrix}
\quad f = (0, 1/6, 1/3, 1/4, 1/6, 1/12, 0)
\]

\[
\Omega_N(f) = \frac{N!}{N_1! \ldots N_k!}
\]
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; $\sum f_j = 1$

$N_j = \text{number of particles in box } \# j$

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

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 \( \nu \);

\[ \hat{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \] (random, "empirical" measure)

What measure shall we observe??

**Informal:** \( \mathbb{P} [\hat{\mu}^N \simeq \mu] \sim e^{-NH_\nu(\mu)} \)

\[ H_\nu(\mu) = \int \rho \log \rho \, d\nu, \quad \rho = \frac{d\mu}{d\nu} \]

**Rigorously:** \( H_\nu = \text{Large Deviation Rate Function of } \hat{\mu}^N \).
Entropy appears in...

- Boltzmann’s theory of gases
- and all kinds of irreversible micro-reversible systems
- Communication and coding theory
- Hydrodynamic limits
- Quantitative central limit theorem
- Nash’s Theorem on regularity for diffusion equations
- Voiculescu’s theory of Von Neumann algebras
- Perelman’s proof of the Poincaré conjecture
- ...
Optimal transport
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 $c(x, y) = d(x, y)^2$

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

$\forall h \quad \int h(x) \pi(dx \, dy) = \int h \, d\mu \quad \int h(y) \pi(dx \, dy) = \int h \, d\nu$
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 $c(x, y) = d(x, y)^2$

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

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

$$(K) \quad \inf_{\pi \in \Pi(\mu, \nu)} \int c(x, y) \pi(dx \, dy)$$
Engineer’s interpretation

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

- **Wasserstein distances** $W_p$: $W_p(\mu_0, \mu_1) = C_p^{1/p}$,

  $C_p = \text{total transport cost with cost } c = d^p$

- **Geodesics**: displace matter along geodesics joining initial to final positions

  $\Rightarrow (\mu_t)_{0 \leq t \leq 1}$
Optimal transport appears in...

- Probability theory, stochastic processes, finance
- Dynamical systems
- Meteorology, fluid mechanics
- Fully nonlinear elliptic equations
- Image processing/interpolation, color engineering
- Urban planning, irrigation, crowd motion modelling
- Study of the cut locus, nonsmooth geometric problems
- Reconstruction of the early Universe
- Functional inequalities...
Unexpected encounter of the third type

(1998) Jordan–Kinderlehrer–Otto: The nonsmooth gradient flow of Boltzmann’s information, in the geometry of $W_2$, is the heat equation!

(1999) Otto–V: Ricci curvature bounds determine convexity properties for Boltzmann’s information along geodesics of $W_2$

Precursors: Marton, McCann

Developed by:
Cordero-Erausquin–McCann–Schmuckenschläger,
Sturm–Von Renesse...

Key cost: square of the distance
The Lazy Gas experiment
\[ S = - \int \rho \log \rho \]
Relation between transport and Ricci

\[ \text{Ric} \geq 0 \]

if and only if

\[ H(\mu_t) = \int \rho_t \log \rho_t \, d\text{vol} \] is a convex function of \( t \)

\[ \rho_t = \frac{d\mu_t}{d\text{vol}} \]

(convexity along geodesics of optimal transport!)
**Definition:** A compact metric-measured space $(X, d, \nu)$ has **Ricci curvature** $\geq 0$ (in weak sense) if

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

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

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

(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 \((X, d, \nu)\) has Ricci curvature \(\geq K\) (in weak sense) if

\[
\forall \mu_0, \mu_1 \in P(X) \exists (\mu_t)_{0 \leq t \leq 1}, \text{ geodesic in } P(X), \text{ s.t.}
\forall t \in [0, 1],
\]

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

- 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:

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

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

where $\pi$ is optimal, $\beta_t(x, y) = \text{reference distortion coeff}$

Two competing choices of reference distortion coefficients

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

$\beta_t(x, y) = \left( \frac{\sin(t\alpha)}{t \sin \alpha} \right)^N \quad \alpha = \sqrt{\frac{K}{N}} \, d(x, y) \quad [\text{CD}^*]$
Consistency
The weak definition coincides with the usual one if the space is smooth (Riemannian manifold)

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

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 $\gamma_x(0) = \nabla \psi(x)$ for some function $\psi$ having some convexity-type properties.

2. The interpolation $\mu_t$ between $\mu_0$ and $\mu_1$ is obtained by stopping the geodesic at time $t$: $\mu_t = (\exp t \nabla \psi) \# \mu_0$

3. Change variables:
$$H(\mu_t) = H(\mu_0) - \int \log \text{Jac}(\exp t \nabla \psi) \, d\mu_0$$
4. $\text{Ric} \geq 0 \implies \frac{d^2}{dt^2} \log \text{Jac}(\exp t\nabla \psi) \leq 0$

... All the rest is “analysis” and approximation...

**Note:**

- The entropy is an “integrated” way to involve the logarithmic Jacobian determinant of the exponential map (Cf. Weak Solutions!)
- With optimal transport only access to gradient velocity fields – rich enough (Hamilton–Jacobi eq.)
- Optimal transport in general is **not smooth** (even on an ellipsoid!) – smoothness is another (beautiful) story
**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

\[ \ddot{D} + \frac{K}{N} D \leq 0 \]

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 \( CD^*(K, N) \) 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 $f_k : \mathcal{X}_k \to \mathcal{X}$ such that

\[
\begin{align*}
|d(f_k(x), f_k(y)) - d_k(x, y)| &\leq \varepsilon_k \to 0 \\
\forall x \in X, \quad d(x, f_k(X_k)) &\leq \varepsilon_k \\
(f_k)_#\nu_k &\rightarrow \nu \quad \text{weakly}
\end{align*}
\]
**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 \(f_k : \mathcal{X}_k \to \mathcal{X}\)

\[
\begin{align*}
|d(f_k(x), f_k(y)) - d_k(x, y)| &\leq \varepsilon_k \to 0 \\
\forall x \in X, \quad d(x, f_k(X_k)) &\leq \varepsilon_k \\
(f_k)_#\nu_k &\to \nu \quad \text{weakly}
\end{align*}
\]

**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 “Ric + $\nabla^2V \geq 0$”:

For any two probability measures $\mu_0$ and $\mu_1$, there is a geodesic $(\mu_t)_{0 \leq t \leq 1}$ in the Wasserstein space $(P(X), W_2)$, 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. $\mu$ and $\nu$,

and satisfies a contraction principle:

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

Conclude that the same property holds true in the limit space, deduce $\text{Ric} + \nabla^2 V \geq 0$. 
**Compatibility (Petrunin 2009)**

If \((X, d)\) is a compact finite-dimensional Alexandrov space with “sectional” curvature \(\geq 0\) then also \((X, d, \text{vol})\) has “Ricci” curvature \(\geq 0\).

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 CD\((K, N)\) spaces.

But weak CD\((K, N)\) spaces are **more general** and include all MGH limits of 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

\( A_0, A_1 \text{ given} \)

\[ \mu_0 := \nu|_{A_0}, \quad \mu_1 := \nu|_{A_1}; \quad (\mu_t)_{0 \leq t \leq 1} \]

\[ \int \rho_{1/2}^{1-1/N} \, d\nu \geq \frac{1}{2}(|A_0|^{1/N} + |A_1|^{1/N}), \]

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 CD(0, ∞),

\[ \lambda_k(M, \nu) \leq C^k \lambda_1(M, \nu) \] for some universal \( C \).

The key is the entropy interpretation and a recursive estimate on the separation: \( \text{Sep}(M, \nu, \alpha_1, \ldots, \alpha_N) := \) maximum min-distance between sets \( A_1, \ldots, A_N \) satisfying \( \nu[A_j] = \alpha_j \), obtained through displacement convexity of \( H \)
Isoperimetric inequalities, concentration

The transport approach gives a grip on measures/sets
Used for concentration inequalities

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

The key is the entropy interpretation and a recursive estimate on the separation: $\text{Sep}(M, \nu, \alpha_1, \ldots, \alpha_N) := \max \min$-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...
- Recently improved by Liu: $C_k = C k^2$,
also for any $\text{CD}(0, \infty)$ finite-dim Alexandrov space
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 \(H_\nu\) in \(P_2\)
- or as \(L^2\) grad flow of Dirichlet form \(\int |\nabla \rho|^2 \, d\nu\)

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
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...)

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

\[ \Rightarrow \text{In any geodesic space, } \partial_t Q_t f + \frac{|\nabla Q_t f|^2}{2} = 0 \text{ (except at countably many times)} \]
How wide is this generalization?

Nonbranching 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
RCD\((K, N)\) Spaces / 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 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.
- Quantitative Splitting Thm (Gigli); sharp inequalities for RCD\(^*\), covariant derivatives, Hessian, curvature...
- A.e. existence of (unique?) tangent spaces, rectifiability (Mondino–Naber)
Recent success

Synthetic proof of the Lévy–Gromov in a nonsmooth (nonbranching) context (Cavalletti–Mondino)

- Adapts to Finsler geometries
- Way simpler, even in the smooth case
- Goes through nonsmooth localization, involving $L^1$-transport
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, D.Paulin...

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

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

Maas-Erbar: Can be made more precise: discrete Ricci curvature of the hypercube is (somehow) $K = 1/(2N)$