Optimal Transport in Geometry, Analysis and Probability

Karl-Theodor Sturm

Universität Bonn

Outline

- **L^2-Wasserstein Space and its Geometry**
- **Gradients and Gradient Flows in the Wasserstein Space**

- **Optimal Transport and Ricci Curvature**
- **Heat Flow on Metric Measure Spaces**

- **Optimal Transport from the Lebesgue Measure to the Poisson Point Process**

L^2-Wasserstein Space

Let (M, d) complete separable metric space, define

$$\mathcal{P}_2(M) \quad = \quad \left\{ \text{prob. meas. } \mu \text{ on } M \text{ with } \int_M d^2(x, x_0)\, \mu(dx) < \infty \right\}$$

and

$$d_W(\mu_0, \mu_1) = \inf_q \left[\int_{M \times M} d^2(x, y)\, d\, q(x, y) \right]^{1/2}$$

where $\inf_q$ over q with $(\pi_1)_* q = \mu_0$, $(\pi_2)_* q = \mu_1$. Then

- $(\mathcal{P}_2(M), d_W)$ is a complete separable metric space.
- $(\mathcal{P}_2(M), d_W)$ is a compact space or a length space or an Alexandrov space[1] with curvature ≥ 0
  if and only if (M, d) is so.

1) Geodesic space with Pythagorean inequality $a^2 + b^2 \geq c^2$

The Problems of Monge and Kantorovich

Monge 1781: Given probab.measures μ, ν, minimize transport costs

$$\int_{\mathbb{R}^n} |x - F(x)|^p \, d\mu(x)$$

among all transport maps $F : \mathbb{R}^n \to \mathbb{R}^n$ with $F_*\mu = \nu$,

i.e. $\mu(F^{-1}(A)) = \nu(A)$ for all $A \subset \mathbb{R}^n$

μ distribution of producers, ν distribution of consumers

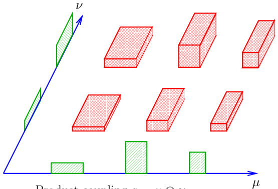
Kantorovich 1942: Minimize

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^p \, d\, q(x, y)$$

among all couplings q of μ and ν,

i.e. $q(A \times \mathbb{R}^n) = \mu(A)$, $q(\mathbb{R}^n \times B) = \nu(B)$ for all $A, B \subset \mathbb{R}^n$

Examples: $q = \mu \otimes \nu$, $q = (Id, F)_*\mu$



Product coupling $q = \mu \otimes \nu$

The Problems of Monge and Kantorovich

Monge 1781: Given probab.measures μ, ν, minimize transport costs

$$\int_{\mathbb{R}^n} |x - F(x)|^p \, d\mu(x)$$

among all transport maps $F : \mathbb{R}^n \to \mathbb{R}^n$ with $F_*\mu = \nu$,

i.e. $\mu(F^{-1}(A)) = \nu(A)$ for all $A \subset \mathbb{R}^n$

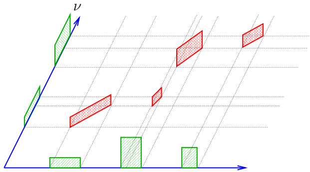
μ distribution of producers, ν distribution of consumers

Kantorovich 1942: Minimize

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^p \, d\, q(x, y)$$

among all couplings q of μ and ν,

i.e. $q(A \times \mathbb{R}^n) = \mu(A)$, $q(\mathbb{R}^n \times B) = \nu(B)$ for all $A, B \subset \mathbb{R}^n$



Optimal coupling q

Examples: $q = \mu \otimes \nu$, $q = (Id, F)_* \mu$

The Problems of Monge and Kantorovich

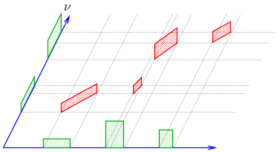
Brenier 1987: "Kantorovich = Monge" if $\mu \ll \mathfrak{L}^n$ and $p = 2$

$$\inf_q \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 \, d\, q(x, y) = \inf_F \int_{\mathbb{R}^n} |x - F(x)|^2 \, d\mu(x)$$

Moreover, $q = (Id, F)_* \mu$, $F = \nabla \varphi$

$\exists$ unique minimizer q,
$\exists$ unique minimizer F,
$\exists$ unique convex $\varphi : \mathbb{R}^n \to \mathbb{R}$ (up to const.)
s.t.

$$\nu = (\nabla \varphi)_* \mu.$$



Optimal coupling q

Example. $n = 1$, $\mu = \mathfrak{L}\big|_{[0,1]}$. Then for all ν on $\mathbb{R}^1$

$$\nu = F_* \mu$$

with F right inverse of distribution function $t \mapsto \nu((-\infty, t])$.

L^2-Wasserstein Space for Riemannian M

(M, g) complete n-dim. Riemannian manifold, m Riemannian volume measure
$c(x, y) = \frac{1}{2} d^2(x, y)$ with Riemannian distance d.

Theorem (McCann 2001). $\forall$ prob. measures μ_0, μ_1 with compact supports and $d\mu_0 \ll dm$:

$\exists$ unique minimizer q, $\exists$ unique minimizer F
$\exists$ unique (up to add. const.) c-convex $\varphi : M \to \mathbb{R} \cup \{\infty\}$ s.t.

$$F(x) = \exp_x (\nabla \varphi(x)) \, .$$

Remarks. (i) On $\mathbb{R}^n$:
φ is c-convex $\iff \varphi_1(x) = \varphi(x) + |x|^2/2$ is convex $\iff \mathrm{Hess}\varphi \geq -I$.
Thus in particular,

$$F(x) = \exp_x (\nabla \varphi(x)) = x + \nabla \varphi(x) = \nabla \varphi_1(x).$$

(ii) If $\varphi : M \to \mathbb{R}$ is c-convex then $\forall t \in (0, 1)$ also $t\varphi$ is c-convex and

$$F_t(x) = \exp_x (t \nabla \varphi(x))$$

is an optimal transport map. For $t > 1$ this may fail.

L^2-Wasserstein Space for Riemannian M

Given $\mu_0, \mu_1 \in \mathcal{P}_2(M)$.
If $\mu_0 \ll m$ then there exists a unique geodesic $(\mu_t)_{0 \le t \le 1}$ connecting them, given as

$$\mu_t := (F_t)_* \mu_0,$$

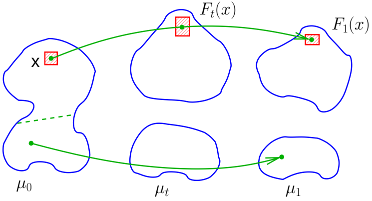
where

$$F_t(x) = \exp_x(t \nabla \varphi(x))$$

with suitable $d^2/2$-convex $\varphi : M \to \mathbb{R}$.

In the case $M = \mathbb{R}^n$ this states that there exists a convex function φ_1 such that

$$F_t(x) = x + t\nabla\varphi(x) = (1-t)x + t\nabla\varphi_1(x).$$



Riemannian Structure of $\mathcal{P}_2(M)$

Tangent space:

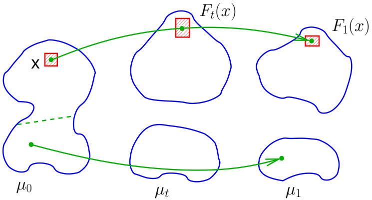
$$T_{\mu_0}\mathcal{P}_2 \quad = \text{closure of} \quad \{\Phi = \nabla\varphi : \ M \to TM, \quad \int_M |\nabla\varphi|^2 d\mu_0 < \infty\}$$

Riemannian tensor:

$$\langle \Phi, \Psi \rangle_{T_{\mu_0}\mathcal{P}_2} \quad = \quad \int_M \Phi \cdot \Psi \, d\mu_0$$

Exponential map:

$$\exp_{\mu_0}(t\,\Phi) \quad = \quad [\exp(t\Phi)]_* \mu_0$$



Riemannian Structure of $\mathcal{P}_2(M)$

Injectivity radius and $d^2/2$-convexity:

- $\nabla\varphi$ is within the injectivity range of $\exp_{\mu_0}$ iff φ is $d^2/2$-convex.
- For each smooth $\varphi : M \to \mathbb{R}$ and for $t > 0$ small enough, $t\varphi$ is $d^2/2$-convex. (That is, Hess $t\varphi \geq -1$ if $M = \mathbb{R}^n$.)

Sectional curvature on $\mathcal{P}_2(M)$

- can be calculated explicitly:

$$\begin{aligned}\mathrm{Sec}_{\mu_0}(\nabla\varphi, \nabla\psi) \quad = \quad & \int_M \mathrm{Sec}_x(\nabla\varphi, \nabla\psi) \cdot \left[|\nabla\varphi|^2 \, |\nabla\psi|^2 - \langle \nabla\varphi, \nabla\psi \rangle^2 \right] d\mu_0 \\ & + 3 \cdot \inf_\eta \int_M \left| \nabla^2\varphi \cdot \nabla\psi - \nabla\eta \right|^2 d\mu_0\end{aligned}$$

for φ, ψ with $\int |\nabla\varphi|^2 d\mu_0 = \int |\nabla\psi|^2 d\mu_0 = 1$, $\int \langle \nabla\varphi, \nabla\psi \rangle d\mu_0 = 0$.

- $\mathrm{Sec} \geq 0$ on M implies $\mathrm{Sec} \geq 0$ on $\mathcal{P}_2(M)$, e.g. $M = \mathbb{R}^n$.
- $\mathrm{Sec} \equiv 0$ on $\mathcal{P}_2(\mathbb{R}^n)$ if and only if $n = 1$.

Riemannian Structure of $\mathcal{P}_2(M)$

For any locally Lipschitz continuous curve $(\mu_t)_{t\in[0,1]} \subset \mathcal{P}_2$ there exists a unique curve $(\Phi_t)_{t\in[0,1]}$ in $T\mathcal{P}_2$

i.e. $\Phi_t \in T_{\mu_t}\mathcal{P}_2$ **for each** $t \in [0,1]$**, in particular,** $\Phi_t(x) \in T_x M$ **for each** $(t,x) \in [0,1] \times M$

satisfying the continuity equation

$$\partial_t \mu_t + \mathrm{div}(\Phi_t \mu_t) = 0$$

in the weak sense, called tangent vector field.

Assume $d\mu(x) \ll dx$ for each $t \in [0,1]$.

Then $(\mu_t)_{t\in[0,1]}$ is a geodesic in $\mathcal{P}_2$ if and only if its tangent vector field is given by $\Phi_t(x) = \nabla\varphi_t(x)$ for some solution $\varphi : [0,1] \times M \to \mathbb{R}$ to the Hamilton-Jacobi equation

$$\frac{\partial}{\partial t}\varphi_t + \frac{1}{2}|\nabla\varphi_t|^2 = 0.$$

Riemannian Structure of $\mathcal{P}_2(M)$

Benamou-Brenier

For any $\mu_0, \mu_1 \in \mathcal{P}_2(M)$

$$d_W(\mu_0, \mu_1) = \inf_{(\mu_t)_t, (\Phi_t)_t} \left(\int_0^1 \|\Phi_t\|^2_{L^2(\mu_t)} \, dt \right)^{1/2},$$

where the infimum is taken over all curves of probability measures $(\mu_t)_{t\in[0,1]}$ on M connecting μ_0 and μ_1 and over all curves of vector fields $(\Phi_t)_{t\in[0,1]}$ on M satisfying weakly the continuity equation

$$\partial_t \mu_t + \mathrm{div}(\Phi_t \mu_t) = 0.$$

Gradients and Gradient Flows in the Wasserstein Space

Calculation of Gradients in the Wasserstein Space

Aim: Determine $\nabla S(\mu) \in T_\mu \mathcal{P}_2(M)$ for 'smooth' $S : \mathcal{P}_2(M) \to \mathbb{R}$.
Easiest example: $S(\mu) = \int_M V\, d\mu$ for given smooth $V : M \to \mathbb{R}$.

Consider geodesic $\mu_t = (F_t)_* \mu_0$ emanating from μ_0 in direction $\Phi \in T_{\mu_0} \mathcal{P}_2(M)$, i.e. $F_t(x) = \exp_x(t\Phi(x)) \approx x + t\Phi(x)$. Then

$$S(\mu_t) = \int V\, d\mu_t = \int V(F_t)\, d\mu_0$$

with $V(F_t) \approx V + t\nabla V \cdot \Phi$. Hence,

$$\frac{\partial}{\partial t} S(\mu_t)\big|_{t=0} = \int \nabla V \cdot \Phi\, d\mu_0$$

and thus

$$\nabla S(\mu_0) = \nabla V \qquad \text{indep. of } \mu_0.$$

More general: $S(\mu) = u\left(\int V_1\, d\mu, \ldots, \int V_N\, d\mu\right)$ 'smooth cylinder function'. Then

$$\nabla S(\mu) = \sum_{i=1}^N \partial_i u(.) \cdot \nabla V_i .$$

Calculation of Gradients in the Wasserstein Space

Main example: Relative entropy
$S(\mu) = \int_M \rho \log \rho \, dm$ for $\mu \ll m$ with density ρ.

Consider geodesic $\mu_t = (F_t)_* \mu_0$ as above with $F_t(x) = \exp_x(t\nabla\varphi(x)) \approx x + t\nabla\varphi(x)$. For t suff. small $\mu_t \ll m$ with density ρ_t satisfying

$$\rho_t(F_t(x)) \cdot \det DF_t(x) = \rho_0(x).$$

(Proof: $\int u(F_t)\rho_0 dm = \int u\rho_t dm = \int u(F_t)\rho_t(F_t) \det DF_t dm$ for all u.)
Hence,

$$S(\mu_t) = \int \rho_t \log \rho_t dm = \int \rho_t(F_t) \log \rho_t(F_t) \det DF_t dm = \int \log \frac{\rho_0}{\det DF_t} \rho_0 dm = S(\mu_0) - \int \log \det DF_t \rho_0 dm.$$

Now $DF_t \approx I + t\nabla^2\varphi$ and $\log \det DF_t \approx \log(1 + t \cdot \operatorname{tr} \nabla^2\varphi) \approx t\Delta\varphi$. Therefore,

$$\frac{\partial}{\partial t} S(\mu_t)\big|_{t=0} = -\int \Delta\varphi \cdot \rho_0 dm = \int \nabla\varphi \cdot \nabla\rho_0 \, dm = \int \nabla\varphi \cdot \nabla \log \rho_0 \cdot \rho_0 \, dm$$

and thus

$$\nabla S(\mu_0) = \nabla \log \rho_0.$$

Analogously: $S(\mu) = \int U(\rho) dm$.

Gradient Flows

The gradient flow

$$\frac{\partial \mu}{\partial t} = -\nabla S(\mu) \qquad \text{on } \mathcal{P}_2(M)$$

for the relative entropy $S(\rho\, dx) = \int \rho \cdot \log \rho \; dx$ is given by $\mu_t(dx) = \rho_t(x)dx$ where ρ solves the heat equation

$$\frac{\partial}{\partial t}\rho = \triangle \rho \qquad \text{on } M.$$

$\mathbb{R}^n$:	**Jordan/Kinderlehrer/Otto '98, Otto '01**
Riemann (M, g):	**Ohta '09, Savare '09, Villani '09, Erbar '09**
Finsler (M, F, m):	**Ohta/Sturm '09**
Wiener space:	**Fang/Shao/Sturm '09**
Heisenberg group:	**Juillet '09**
Alexandrov spaces:	**Gigli/Kuwada/Ohta '10**
Metric meas. spaces:	**Ambrosio/Gigli/Savare '11**
Discrete spaces:	**Maas '11**

Gradient Flows

Consider $\frac{\partial \mu_t}{\partial t} = -\nabla S(\mu_t)$ and recall $\nabla S(\mu_t) = \nabla \log \rho_t$. Thus

$$\mu_{t+h} \approx (F_{t,t+h})_* \mu_t \quad \text{with} \quad F_{t,t+h}(x) = \exp_x(h \frac{\partial \mu_t}{\partial t}) \approx x - h\nabla \log \rho_t.$$

Since $\int u d\mu_{t+h} = \int u(F_{t,t+h}) d\mu_t$ for all $u : M \to \mathbb{R}$ we obtain

$$\begin{aligned} \frac{\partial}{\partial t} \int u \rho_t \, dm &= \frac{\partial}{\partial h}\Big|_{h=0} \int u d\mu_{t+h} = \frac{\partial}{\partial h}\Big|_{h=0} \int u(F_{t,t+h}) d\mu_t \\ &= -\int \nabla u \cdot \nabla \log \rho_t \, d\mu_t = \int u \, \Delta \rho_t \, dm. \end{aligned}$$

In other words,

$$\frac{\partial}{\partial t} \rho_t = \Delta \rho_t.$$

Gradient Flows

Consider

$$S(\nu) = \frac{1}{s-1}\int \rho^s \, dx + \int V d\nu + \int\int W d\nu \, d\nu$$

for $d\nu = \rho \, dx + d\nu^{sing}$.
Here $s > 0$ real, $V : \mathbb{R}^n \rightarrow \mathbb{R}$ some external potential and
$W : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ some interaction potential.

Theorem. (Jordan/Kinderlehrer/Otto '98, Otto '01, Villani '03, Ambrosio/Gigli/Savare '05, . . .)

The gradient flow $\frac{\partial \nu}{\partial t} = -\nabla S(\nu)$ on $\mathcal{P}_2(\mathbb{R}^n)$ is given by $\nu_t(dx) = \rho_t(x)dx$ where ρ solves the nonlinear PDE

$$\frac{\partial}{\partial t}\rho = \triangle(\rho^s) + \nabla(\rho \cdot \nabla V) - \nabla(\rho \cdot \int(\nabla W \rho))$$

This includes porous medium equation, fast diffusion, Fokker-Planck, McKean-Vlasov.

Other examples:

quantum-drift diffusion (Fisher information), Ginzburg-Landau dynamics (squared H^{-1}-norm), p-Laplacian.

Equilibration — Functional Inequalities (Otto-Villani)

Convexity properties of S on $\mathcal{P}_2(M)$ imply equilibration estimates for the gradient flow $(t, \nu_0) \mapsto \nu_t$ on $\mathcal{P}_2(M)$

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

$$\Downarrow$$

$$|\nabla S|^2 \geq 2\, K \cdot S$$

$$\Downarrow$$

$$S \geq K/2 \cdot d_W(\cdot, \nu_\infty)^2$$

$$\Downarrow$$

$$d_W(\nu_t, \nu_\infty) \leq C \cdot e^{-K\, t}$$

Equilibration — Functional Inequalities (Otto-Villani)

Convexity properties of S on $\mathcal{P}_2(M)$ imply equilibration estimates for the gradient flow $(t, \nu_0) \mapsto \nu_t$ on $\mathcal{P}_2(M)$

$\mathrm{Hess}\, S \geq K$	**"Ricci curvature bound"**
$\Downarrow$	
$\lvert\nabla S\rvert^2 \geq 2\, K \cdot S$	**"log. Sobolev inequality"**
$\Downarrow$	
$S \geq K/2 \cdot d_W(\cdot, \nu_\infty)^2$	**"Talagrand inequality"**
$\Downarrow$	
$d_W(\nu_t, \nu_\infty) \leq C \cdot e^{-K\, t}$	**"Exponential Decay"**

Example: McKean-Vlasov

Estimating rate of convergence to equilibrium for McKean-Vlasov equation on $\mathbb{R}^d$

$$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u - \nabla(u \cdot \nabla V * u)$$

with symmetric, uniformly continuous interaction potential $V: \mathbb{R}^d \to \mathbb{R}$.

I. Approximation by interacting particle systems (+ factorization):

$$dX_t^i = \frac{1}{N}\sum_{i=1}^{N} \nabla V(X_t^i - X_t^j)dt + dB_t^i$$

- Propagation of chaos, hydrodynamic limit: $\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^i} \to \mu_t(dx) = u(t,x)dx$
- Bakry-Émery curvature condition applied to $d\overline{X}_t = \nabla\overline{V}(\overline{X}_t)\,dt + d\overline{B}_t$, where $\overline{X}_t = (X_t^1, \ldots, X_t^N)$:

$\mathrm{Hess}\, V \geq K > 0 \quad \Rightarrow \quad \mathrm{Hess}\,\overline{V} \geq K \quad \Rightarrow \quad \overline{X}_t$ converges with rate K $\Rightarrow \quad u_t$ converges with rate K

Example: McKean-Vlasov

Estimating rate of convergence to equilibrium for McKean-Vlasov equation on $\mathbb{R}^d$

$$\frac{\partial u}{\partial t} = \frac{1}{2} \triangle u - \nabla (u \cdot \nabla V * u)$$

with symmetric, uniformly continuous interaction potential $V : \mathbb{R}^d \to \mathbb{R}$.

II. Approach via mass transportation

$\mu_t(dx) = u(t,x)dx$ is gradient flow in $\mathcal{P}_2(\mathbb{R}^d)$ of

$$S(\mu) = \mathrm{Ent}(\mu|dx) + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V(x-y)\mu(dx)\mu(dy).$$

$\mathrm{Hess} V \geq K > 0$ on $\mathbb{R}^d$ $\Rightarrow$ $\mathrm{Hess} S \geq K$ on $\mathcal{P}_2(\mathbb{R}^d)$ $\Rightarrow$ u_t converges with rate K.

Benachour et al., Carillo/McCann/Villani, Malrieu

Optimal Transport and Ricci Curvature

OT, Entropy, and Ricci Curvature

M complete Riemannian manifold, m Riemannian volume measure

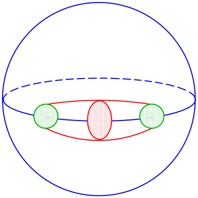
Theorem. (Otto '01, Otto/Villani '00, Cordero/McCann/Schmuckenschläger '01, vRenesse/Sturm '05)

Let $\mathrm{Ent}(\mu) = \int \rho \cdot \log \rho \; dm$ with $\rho = \frac{d\mu}{dm}$. Then

$$\mathrm{Hess\,Ent} \geq K \quad \Leftrightarrow \quad \mathrm{Ric}_M \geq K$$

The **proof** depends on the following estimate for the logarithmic determinant $y_t := \log\det dF_t$ of the Jacobian of the transport map:

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



This inequality is sharp. It describes the effect of curvature on optimal transportation.

OT, Entropy, and Ricci Curvature

Consider $t \mapsto (F_t)_* \mu_0 = \mu_t = \rho_t m$ geodesic in $\mathcal{P}_2(M)$.

$$\begin{aligned} \mathrm{Ent}(\mu_t) &= \int \rho_t \log \rho_t \, dm \\ &\overset{(i)}{=} \int \rho_t(F_t) \cdot \log \rho_t(F_t) \det dF_t \, dm \\ &\overset{(ii)}{=} \int \rho_0 \cdot (\log \rho_0 - y_t) \, dm \\ &= \mathrm{Ent}(\mu_0) - \int y_t \, d\mu_0 \end{aligned}$$

$$\Downarrow$$

$$\partial_t^2 \mathrm{Ent}(\mu_t) = - \int \ddot{y}_t \, d\,\mu_0 \overset{(iii)}{\geq} K \cdot \int |\dot{F}_t|^2 \, d\,\mu_0 = K \cdot d_W^2(\mu_0, \mu_1)$$

(i) Change of variables **(ii) Transport property** $\rho_t(F_t) \cdot \det dF_t = \rho_0$

(iii) Basic estimate for $y_t = \log \det dF_t$.

OT, Entropy, and Ricci Curvature

- Gradient flow for $S(\rho) = \int \rho \cdot \log \rho \, dx$ satisfies $\frac{\partial}{\partial t}\rho = \triangle \rho$ and

$$\mathrm{Hess}\, S \; \geq K \qquad \Leftrightarrow \qquad \mathrm{Ric}_M \geq K$$

$\rightsquigarrow$ **Ricci bounds for metric measure spaces**

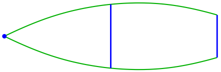
$\rightsquigarrow$ **logarithmic Sobolev inequality, concentration of measure**

- Grad. flow for $S(\rho) = \frac{1}{m-1} \int_M \rho^m(x) \, dx$ satisfies $\frac{\partial}{\partial t}\rho = \triangle(\rho^m)$ and

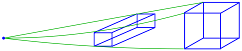
$$\mathrm{Hess}\, S \; \geq 0 \quad \Leftrightarrow \quad \begin{cases} m & \geq \; 1 - \frac{1}{n} \\ \mathrm{Ric}_M & \geq \; 0 \end{cases}$$

$\rightsquigarrow$ **Curvature-Dimension condition CD(K,N) for mms**

$\rightsquigarrow$ **Sobolev inequality, Bishop-Gromov volume growth estimates**



sec $\geq \mathbf{0}$ $\iff$ dist concave



ric $\geq \mathbf{0}$ $\iff$ $\mathrm{vol}^{1/n}$ concave

Ricci Bounds for Metric Measure Spaces (M, d, m)

(M, d) complete separable metric space, m locally finite measure

Definition. $\mathrm{Ric}(M, \mathsf{d}, m) \geq K$ or $CD(K, \infty)$

$\Longleftrightarrow \quad \forall \mu_0, \mu_1 \in \mathcal{P}_2(M): \ \exists$ geodesic $(\mu_t)_t$ s.t. $\forall t \in [0,1]$:

$$\mathrm{Ent}(\mu_t|m) \leq (1-t)\,\mathrm{Ent}(\mu_0|m) + t\,\mathrm{Ent}(\mu_1|m) - \frac{K}{2}\,t(1-t)\,\mathrm{d}_{\mathrm{w}}^2(\mu_0, \mu_1)$$

Recall
$$\mathrm{Ent}(\nu|m) = \begin{cases} \int_M \rho \log \rho \, dm & , \text{ if } \nu = \rho \cdot m \\ +\infty & , \text{ if } \nu \not\ll m \end{cases}$$

The Condition $CD(K, N)$

Definition. A metric measure space (M, d, m) satisfies the **Curvature-Dimension Condition** $CD(K, N)$ for $K, N \in \mathbb{R}$ iff the Rényi type entropy

$$S_N(\mu) = -\int_M \rho^{1-1/N} dm$$

satisfies

$$\partial_t^2 S_N(\mu_t) \;\geq\; -\frac{K}{N} S_N(\mu_t)$$

in integrated form along geodesics $(\mu_t)_t$.

That is, for instance,

$$CD(0, N) \quad \Longleftrightarrow \quad \forall \mu_0, \mu_1 \in \mathcal{P}_2(M): \; \exists \text{ geodesic } (\mu_t)_t \text{ s.t. } \forall t \in [0,1]: \\ S_N(\mu_t) \;\leq\; (1-t)\, S_N(\mu_0) + t\, S_N(\mu_1)$$

The Condition $CD(K,N)$

Riemannian manifolds:

$$CD(K,N) \quad \Longleftrightarrow \quad \mathrm{Ric}_M \geq K \quad \text{and} \quad \dim_M \leq N$$

Weighted Riemannian spaces (M,d,m) with $dm = e^{-V} dvol$:
$CD(K,N) \quad \Longleftrightarrow \quad n = \dim_M \leq N$ and

$$\mathrm{Ric}_M + \mathrm{Hess}\, V - \frac{1}{N-n} DV \otimes DV \;\geq\; K$$

Further examples: Alexandrov spaces, Finsler manifolds (e.g. Banach spaces), Wiener space ($K = 1, N = \infty$).

Constructions: Products, cones, suspensions.

The Condition $CD(K, N)$

Theorem. Assume $m(M) = 1$.

$CD(K, N)$ with $K > 0$ and $N \leq \infty$ implies

- Logarithmic Sobolev Inequality
- Talagrand Inequality
- Concentration of Measure
- Poincaré / Lichnerowicz Inequality: for all functions f with $\int_M f\, dm = 0$

$$K \frac{N}{N-1} \cdot \int_M f^2 dm \leq \int_M |\nabla f|^2 dm.$$

The Condition $CD(K,N)$

Theorem. $CD(K,N)$ with $N<\infty$ implies

$$\frac{s(r)}{s(R)} \geq \frac{\sin\left(\sqrt{\frac{K}{N-1}}\,r\right)^{N-1}}{\sin\left(\sqrt{\frac{K}{N-1}}\,R\right)^{N-1}} \qquad \text{for } s(r)=\frac{\partial}{\partial r}m(B_r(x_0))$$

Bishop-Gromov Volume Growth Estimate

Corollary. $CD(K,N)$ with $K>0$ and $N<\infty$ implies

$$\mathrm{diam}(M) \leq \sqrt{\frac{N-1}{K}}\cdot\pi$$

Bonnet-Myers Diameter Bound

Further related inequalities with sharp constants:

Brunn-Minkowski, Prekopa-Leindler, Borell-Brascamp-Lieb

The Condition $CD(K, N)$

Theorem.

The curvature-dimension condition is **stable** under convergence.

Theorem.

For all $K, N, L \in \mathbb{R}$ the space of all (M, d, m) with $CD(K, N)$ and with diameter $\leq L$ is **compact**.

- St.: Acta Math. **196** (2006)
- Lott, Villani: Annals of Math. **169** (2009)

The L^2-Transportation Metric $\mathbb{D}$

$$\mathbb{D}((M_1, d_1, m_1), (M_2, d_2, m_2))^2 = \inf_{d,m} \int_{M_1 \times M_2} d^2(x, y) d\, m(x, y)$$

where the inf is taken over all couplings d of d_1 and d_2 and over all couplings m of m_1 and m_2.

A measure m on the product space $M_1 \times M_2$ is a coupling of m_1 and m_2 iff

$$m(A_1 \times M_2) = m_1(A_1), \quad m(M \times A_2) = m_2(A_2)$$

for all measurable sets $A_1 \subset M_1, A_2 \subset M_2$.

A pseudo metric d on the disjoint union $M_1 \sqcup M_2$ is a coupling of d_1 and d_2 iff

$$d(x, y) = d_1(x, y), \quad d(x', y') = d_2(x', y')$$

for all $x, y \in \mathrm{supp}[m_1] \subset M_1$ and all $x', y' \in \mathrm{supp}[m_2] \subset M_2$.

The L^2-Transportation Metric $\mathbb{D}$

For each pair of normalized mm-spaces (M_1, d_1, m_1) and (M_2, d_2, m_2) there exists an optimal pair of couplings d and m.

$\mathbb{D}$ is a complete separable length metric on the space of isomorphism classes of normalized metric measure spaces.

Measured GH-conv. $\Longrightarrow$ Gromov's $\square$-conv. $\Longleftrightarrow$ $\mathbb{D}$-conv.

The L^2-Transportation Metric $\mathbb{D}$

Given (optimal) coupling q of m_1, m_2 put

$q_1(x, dy) =$ disintegration of $q(dx, dy)$ w.r.t. $m_1(dx)$
$q_2(y, dx) =$ disintegration of $q(dx, dy)$ w.r.t. $m_2(dy)$.

Define map

$$q_2 : \mathcal{P}_{ac}(M_1) \to \mathcal{P}_{ac}(M_2), \ \rho_1 m_1 \mapsto \rho_2 m_2$$

by

$$\rho_2(y) = \int_{M_1} \rho_1(x)\, q_2(y, dx).$$

Then

$$\mathrm{Ent}(\rho_2 m_2 \mid m_2) \ \leq \ \mathrm{Ent}(\rho_1 m_1 \mid m_1).$$

$\rightsquigarrow$ proof of stability result

Heat Flow on Metric Measure Spaces

Heat Flow on Metric Measure Spaces (M, d, m)

Heat equation on M

- either as gradient flow on $L^2(M, m)$ for the **energy**

$$\mathcal{E}(u) = \frac{1}{2} \int_M |\nabla u|^2 \, dm$$

(with "$|\nabla u|$" local Lipschitz constant or minimal upper gradient or Finsler norm or ...)

- or as gradient flow on $\mathcal{P}_2(M)$ for the **relative entropy**

$$\mathrm{Ent}(u) = \int_M u \log u \, dm.$$

Heat Flow on Metric Measure Spaces (M, d, m)

Theorem (Ambrosio/Gigli/Savare '11+).

For arbitrary metric measure spaces (M, d, m) satisfying $CD(K, \infty)$ both approaches coincide.

Corollary. Stability of heat flow under convergence $M_n \to M$.

If heat flow is linear ('Riemannian mms'):

- L^2-Wasserstein contraction

$$d_W(p_t\mu, p_t\nu) \leq e^{-Kt}\, d_W(\mu, \nu)$$

- Bakry-Émery gradient estimate

$$\nabla|p_t u|^2(x) \leq e^{-Kt} \cdot p_t\left(|\nabla u|^2\right)(x)$$

- Coupling of Brownian motions, Lipschitz cont. of harmonic functions

Heat Flow on Metric Measure Spaces (M, d, m)

Particular Cases – Previous Results

Alexandrov spaces

- Both approaches to heat flow coincide
  [Gigli/Kuwada/Ohta '10]
- Bochner inequality, Li-Yau estimates
  [Zhang/Zhu, Qian/Z/Z '10-'12+]

Finsler spaces

- Both approaches coincide $\longrightarrow$ nonlinear heat equation
- L^2-contraction, Bakry-Émery, Bochner inequality, Li-Yau estimates
- No exponential growth rate for L^2-Wasserstein distance

[Ohta/St. '09-'11]

Wiener Space

$M = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, $m =$ Wiener measure, $d =$ Cameron-Martin distance

$$d(x,y) = \left(\int_0^\infty |\dot{x}(t) - \dot{y}(t)|^2 \, dt \right)^{1/2}$$

Transport cost / concentration inequalities

Talagrand, Ledoux, Wang, Fang, Shao, . . . (1996, . . .)

Existence & uniqueness of optimal transport map between m and $\rho\, m$

Feyel/Ustunel (2004)

Gradient flow for the relative entropy $\mathrm{Ent}(.|m)$ on $\mathcal{P}_2(M,d)$

$=$ Ornstein-Uhlenbeck semigroup on M.

Fang/Shao/St.: PTRF (2009)

Discrete Spaces

Let X be a finite space and $K : X \times X \to \mathbb{R}_+$ a Markov kernel, i.e

$$\sum_{y \in X} K(x, y) \; = \; 1 \quad \forall x \in X \; .$$

Assume that K is irreducible and reversible with unique steady state π :

$$\pi(y) \; = \; \sum_{x \in X} \pi(x) K(x, y) \quad \forall y \in X \; .$$

Set $\mathcal{P}(X) = \left\{ \rho : X \to \mathbb{R}_+ \mid \sum_{x \in X} \rho(x)\pi(x) = 1 \right\}$. Given $\rho \in \mathcal{P}(X)$ the entropy is defined by

$$\mathrm{Ent}(\rho) \; = \; \sum_{x \in X} \rho(x) \log \rho(x) \pi(x) \; .$$

The continuous-time Markov semigroup ("heat flow") associated with K is given by $P(t) = e^{t(K - Id)}$.

Discrete Spaces

Given $\rho \in \mathcal{P}(X)$ set $\rho(x,y) = \theta(\rho(x), \rho(y))$, where $\theta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is the logarithmic mean $\theta(s,t) = \frac{s-t}{\log s - \log t}$.

Definition

For $\rho_0, \rho_1 \in \mathcal{P}(X)$ define

$$\mathcal{W}^2(\rho_0, \rho_1) = \inf_{\rho,\psi} \int_0^1 \sum_{x,y \in X} (\psi_t(x) - \psi_t(y))^2 \rho_t(x,y) K(x,y) \pi(x)$$

where $\rho : [0,1] \to \mathcal{P}(X), \psi : [0,1] \to \mathbb{R}^X$ are continuous resp. measurable and satisfy the continuity equation

$$\begin{cases} \frac{d}{dt}\rho_t(x) + \sum_{y \in X} (\psi_t(y) - \psi_t(x)) \rho_t(x,y) K(x,y) = 0 \quad \forall x \in X \ , \\ \rho(0) = \rho_0, \ \rho(1) = \rho_1 \ . \end{cases}$$

Discrete Spaces

Theorem [Maas '11]

$(\mathcal{P}(X), \mathcal{W})$ is a complete metric space. Every pair $\rho_0, \rho_1 \in \mathcal{P}(X)$ is connected by a unique constant speed geodesic, i.e. a curve $\rho : [0,1] \to \mathcal{P}(X)$ satisfying

$$\mathcal{W}(\rho_s, \rho_t) \;=\; |s-t| \mathcal{W}(\rho_0, \rho_1) \quad \forall s, t \in [0,1] \; .$$

The continuous-time Markov chain $P(t) = e^{t(K - Id)}$ evolves as the gradient flow of the entropy w.r.t. $\mathcal{W}$.

Discrete Spaces

Let $X = \mathcal{Q}^n = \{0,1\}^n$ equipped with the usual graph structure and the uniform probability measure π. Let K be the transition probability of simple random walk on $\mathcal{Q}^n$, i.e

$$K(x,y) = \begin{cases} \frac{1}{n}, & x \sim y \\ 0, & \text{else} . \end{cases}$$

Theorem [Erbar-Maas '11]

$(\mathcal{Q}^n, K)$ has Ricci curvature bounded below by $\frac{2}{n}$.

Corollary

The uniform measure π on $\mathcal{Q}^n$ satisfies the modified log-Sobolev inequality:

$$\mathrm{Ent}(\rho|\pi) \ \leq \ \frac{n}{4}\mathcal{I}(\rho|\pi) \ = \ \frac{1}{8}\sum_{x,y\sim x}(\rho_x - \rho_y)(\log\rho_x - \log\rho_y)\,\pi_x.$$

Ongoing Work, Outlook

- "Otto calculus" ("Otto-Villani", "Lott-Sturm-Villani") for Lévy processes [Erbar]:
  semigroups as gradient flows of the entropy in modified L^2-Wasserstein space

- Detailed properties of heat flow and Brownian motion on metric measure spaces with $CD(K, N)$ [Gigli et al.]:
  Laplacian comparison, Bochner inequality ("Bakry-Émery"), differential Harnack inequality ("Li-Yau")

- Gradient flows on the space $\mathbb{X}$ of all metric measure spaces [Sturm]:
  nonnegative curvature on $\mathbb{X}$, semiconvex functions.