Heat flow on Alexandrov spaces

Kazumasa Kuwada

(Ochanomizu University)

joint work with N. Gigli (Nice)

and S. Ohta (Kyoto)

1. Introduction

(X, d): compact Alexandrov space of curv. $\geq k$,

$\partial X = \emptyset$

(geodesic metric sp. of sect. curv. $\geq k$)

$n := \dim_H X \in \mathbb{N}$

$m := \mathcal{H}^n$: Hausdorff measure

(X, d): compact Alexandrov space of curv. $\geq k$,

$$\partial X = \emptyset$$

(geodesic metric sp. of sect. curv. $\geq k$)

$n := \dim_H X \in \mathbb{N}$

$m := \mathcal{H}^n$: Hausdorff measure

- It naturally appears in geometry
- It admits singularity of "curv. $= \infty$"
- Set of singular points can be dense
- Usual differential calculus does not work

Two different ways to define a "heat distribution" on a metric measure space (X, d, m)

(1) Gradient flow of Dirichlet energy functional on L^2-sp. of functions (Dirichlet form)

(2) Gradient flow of relative entropy functional on a sp. of probability measures (Otto calculus)

Two different ways to define a "heat distribution" on a metric measure space (X, d, m)

(1) Gradient flow of Dirichlet energy functional
on L^2-sp. of functions (Dirichlet form)

(2) Gradient flow of relative entropy functional
on a sp. of probability measures (Otto calculus)

——— ∘ ——— ∘ ———

On Riem. mfd, based on PDE theory,

(1) $\rightsquigarrow$ the sol. to the heat eq. $\leftsquigarrow$ (2)

Two different ways to define a "heat distribution" on a metric measure space (X, d, m)

(1) Gradient flow of Dirichlet energy functional on L^2-sp. of functions (Dirichlet form)

(2) Gradient flow of relative entropy functional on a sp. of probability measures (Otto calculus)

——— ○ ——— ○ ———

On Riem. mfd, based on PDE theory,

$$(1) \rightsquigarrow \text{the sol. to the heat eq.} \leftsquigarrow (2)$$

$$\Rightarrow \boxed{(1) = (2)}$$ [Erbar '10]

Q. What happens when X is an Alexandrov space?

Q. What happens when X is an Alexandrov space?

Ans. (1) & (2) coincide on (X, d, m)

Q. What happens when X is an Alexandrov space?

Ans. (1) & (2) coincide on (X, d, m)

⋆ Our argument does not rely on any PDE theory

$$\Downarrow$$

We can combine properties of (1) and (2) in studying heat distributions.

$$\Downarrow$$

We can combine properties of (1) and (2) in studying heat distributions.

$$\Downarrow$$

Theorem [Gigli-K.-Ohta]

The heat kernel $p_t(x, \cdot)$ is Lipschitz continuous

(It improves the known Hölder continuity coming from the theory of Dirichlet forms)

2. Framework

2.1. Dirichlet energy and its gradient flow

[Kuwae, Machigashira & Shioya '01]

$\exists(\mathcal{E}, W^{1,2}(X))$: (str. local, reg.) Dirichlet form

$$\mathcal{E}(u,u) := \int_X \langle \nabla u, \nabla u \rangle \, dm$$

$$(u \in W^{1,2}(X))$$

$(\mathcal{E}, W^{1,2}(X)) \leftrightarrow (\Delta, \mathcal{D}(\Delta))$: generator

$\leftrightarrow T_t = \mathrm{e}^{t\Delta}$: semigroup

Properties

- $\mathrm{Lip}(X) \subset W^{1,2}(X)$ dense.

  Moreover, for $f \in \mathrm{Lip}(X)$,

$$|\nabla_d f| \left(= \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x,y)} \right)$$
$$= |\nabla f| \quad m\text{-a.e.}$$

- $\exists$(Hölder) conti. heat kernel $p_t(x, y)$:

$$T_t f(x) = \int_X p_t(x,y) f(y) m(dy)$$

2.2. Gradient flow of relative entropy on $(\mathcal{P}_2(X), d_2^W)$

L^2-Wasserstein distance

For $\mu_0, \mu_1 \in \mathcal{P}(X)$,

$$d_2^W(\mu_0, \mu_1) := \inf \left\{ \|d\|_{L^2(\pi)} \;\middle|\; \begin{array}{l} \pi\text{: coupling of} \\ \mu_0 \text{ and } \mu_1 \end{array} \right\}$$

⋆ $(\mathcal{P}(X), d_2^W)$: cpt. geodesic metric sp.,

compatible with the weak conv.

Detailed description of the L^2-Wasserstein distance

$\Gamma = \{\gamma : [0,1] \to X$ const. speed min. geod.$\}$

$\mathrm{e}_t : \Gamma \to X$, $\mathrm{e}_t(\gamma) := \gamma(t)$

$\star$ $\exists\, \Pi \in \mathcal{P}(\Gamma)$ s.t.

- $e_0^{\#}\Pi = \mu_0$, $e_1^{\#}\Pi = \mu_1$
- $d_2^W(\mathrm{e}_t^{\#}\Pi, \mathrm{e}_s^{\#}\Pi) = (t - s)\, d_2^W(\mu_0, \mu_1)$

[Lott & Villani '09]

$\star$ When $\nu \ll m$, $\exists F : X \to \Gamma$ s.t. $\Pi = F^{\#}\mu$,

([Bertrand '08] in our framework)

Relative entropy

$$\mathbf{Ent}(\mu) := \begin{cases} \int_X \rho \log \rho \, dm & \text{if } d\mu = \rho dm \\ \infty & \text{otherwise} \end{cases}$$

Heuristics:

Why does gradient flow of Ent on $(\mathcal{P}(X), d_2^W)$ provide the sol. to the heat eq.?

Key observation

The formal Riemannian structure is compatible with the L^2-Wasserstein distance d_2^W

For $\mu_t := \Phi_t^{\#}\mu$ and $\nu_t := \Psi_t^{\#}\mu$,

$$\langle \dot{\mu}_0, \dot{\nu}_0 \rangle := \int_X \langle \partial_t \Phi_0, \partial_t \Psi_0 \rangle d\mu$$

When $\mu_t = (e_t \circ F)^{\#}\mu = e_t^{\#}\Pi$: min. geod.,

$$\begin{aligned}\int_0^1 |\dot{\mu}_t|^2 dt &= \int_0^1 \int_X |\partial_t (e_t \circ F)|^2 \, d\mu dt \\ &= \int_\Gamma \int_0^1 |\dot{\gamma}_t|^2 \, dt \Pi(d\gamma) \\ &= \int_\Gamma d(e_0(\gamma), e_1(\gamma))^2 \Pi(d\gamma) \\ &= d_2^W(\mu_0, \mu_1)^2\end{aligned}$$

For $\mu_t = \Phi_t^{\#}\mu_0 = \rho_t m$,

$$\partial_t \mathrm{Ent}(\mu_t)|_{t=0} = \partial_t \int_X \log \rho_t \, d\mu_t \Big|_{t=0}$$

$$= \partial_t \int_X \log \rho_t(\Phi_t) \, d\mu_0 \Big|_{t=0}$$

$$= \int_X (\partial_t \rho)_0 dm + \int_X \langle \frac{\nabla \rho_0}{\rho_0}, \partial_t \Phi_0 \rangle d\mu_0$$

$$\Rightarrow \boxed{\dot{\mu}_t = -\nabla \mathrm{Ent}(\mu_t) \text{ iff } \partial_t \Phi_t = -\frac{\nabla \rho_t}{\rho_t}}$$

$\Rightarrow$ When $\dot{\mu}_t = -\nabla \mathrm{Ent}(\mu_t)$,

$$\begin{aligned}
\partial_t \int_X g \, d\mu_t \Big|_{t=0} &= \partial_t \int_X g(\Phi_t) \, d\mu_0 \Big|_{t=0} \\
&= \int_X \langle \nabla g, \partial_t \Phi_0 \rangle d\mu_0 \\
&= -\int_X \langle \nabla g, \nabla \rho_0 \rangle dm \\
&= \int_X \Delta g \, d\mu_0
\end{aligned}$$

$\therefore$ μ_t solves the (weak) heat equation.

Definition of the grad. flow $(\mu_t)_{t\geq 0}$

$(\mu_t)_{t\geq 0}$: abs. conti., $\mathrm{Ent}(\mu_t) < \infty$,

$$\mathrm{Ent}(\mu_t) - \mathrm{Ent}(\mu_s) = -\frac{1}{2}\int_s^t |\dot{\mu}_r|^2 dr - \frac{1}{2}\int_s^t |\nabla_- \mathrm{Ent}(\mu_r)|^2 dr$$

for $0 \leq s \leq t$, where

$$|\dot{\mu}_r| := \limsup_{h\downarrow 0} \frac{1}{h} d_2^W(\mu_{r+h}, \mu_r)$$

$$|\nabla_- \mathrm{Ent}(\mu)| := \limsup_{\nu\to\mu} \frac{[\mathrm{Ent}(\mu) - \mathrm{Ent}(\nu)]_+}{d_2^W(\mu,\nu)}$$

Heuristics:

Why does this definition work?

$$\mathbf{Ent}(\mu_t) - \mathbf{Ent}(\mu_s)$$

$$\text{“}=\text{”} \int_s^t \langle \dot{\mu}_r, \nabla \mathbf{Ent}(\mu_r) \rangle dr$$

$$\geq -\frac{1}{2} \int_s^t |\dot{\mu}_r|^2 dr - \frac{1}{2} \int_s^t |\nabla \mathbf{Ent}|^2 (\mu_r) dr$$

$$\left(\because \langle u, v \rangle \geq -\frac{1}{2} (\langle u, u \rangle + \langle v, v \rangle) \right)$$

and “$=$” holds iff $\dot{\mu}_r = -\nabla \mathbf{Ent}(\mu_r)$

The condition $\mathsf{CD}(K,\infty)$

For $\forall (\nu_t)_{t\in[0,1]}$: d_2^W-min. geod.,

$$\mathrm{Ent}(\nu_\lambda) \le (1-\lambda)\mathrm{Ent}(\nu_0) + \lambda \mathrm{Ent}(\nu_1) - \frac{K}{2}\lambda(1-\lambda) d_2^W(\nu_0,\nu_1)^2$$

(K-convexity of Ent w.r.t. d_2^W)

The condition CD(K,∞)

For $\forall(\nu_t)_{t\in[0,1]}$: d_2^W-min. geod.,

$$\mathrm{Ent}(\nu_\lambda) \le (1-\lambda)\mathrm{Ent}(\nu_0) + \lambda\mathrm{Ent}(\nu_1) - \frac{K}{2}\lambda(1-\lambda)d_2^W(\nu_0,\nu_1)^2$$

(K-convexity of Ent w.r.t. d_2^W)

- [von Renesse & Sturm '05]
  CD$(K,\infty) \Leftrightarrow \mathrm{Ric} \ge K$ if X: Riem. mfd

- [Petrunin]
  (X,d,m) satisfies CD$((n-1)k,\infty)$

Existence and uniqueness of gradient flow

Under CD(K,∞), $\exists!$ grad. flow of Ent for $\forall$ initial $\mu \in \mathcal{P}(X)$ with $\mathrm{Ent}(\mu) < \infty$

[Ambrosio, Gigli & Savaré '05, Ohta '09, Gigli '10]

Existence and uniqueness of gradient flow

Under CD(K, ∞), $\exists!$ grad. flow of Ent

for $\forall$ initial $\mu \in \mathcal{P}(X)$ with $\mathrm{Ent}(\mu) < \infty$

[Ambrosio, Gigli & Savaré '05, Ohta '09, Gigli '10]

L^2-Wasserstein contraction

For grad. flows μ_t and $\tilde{\mu}_t$,

$$d_2^W(\mu_t, \tilde{\mu}_t) \le \mathrm{e}^{-Kt} d_2^W(\mu_0, \tilde{\mu}_0)$$

[Savaré '07, Ohta '09, Gigli & Ohta '10]

$\Rightarrow$ Grad. flow for any initial $\mu \in \mathcal{P}(X)$

3. Main result and sketch of the proof

Theorem 1 [Gigli-K.-Ohta]

For any $\mu \in \mathcal{P}(X)$,

$T_t\mu$ is a gradient flow of Ent on $(\mathcal{P}(X), d_2^W)$

Suppose $\mathrm{Ent}(\mu) < \infty$. $\mu_t := T_t\mu$, $\rho_t := \dfrac{d\mu_t}{dm}$.

Goal

$$\mathrm{Ent}(\mu_t) - \mathrm{Ent}(\mu_s) = -\frac{1}{2}\int_s^t |\dot{\mu}_r|^2 dr - \frac{1}{2}\int_s^t |\nabla_- \mathrm{Ent}(\mu_r)|^2 dr$$

- "$\geq$" is always true
- Sufficient to show: for a.e.t,

$$\partial_t \mathrm{Ent}(\mu_t) + \frac{1}{2}|\dot{\mu}_t|^2 + \frac{1}{2}|\nabla_- \mathrm{Ent}(\mu_t)|^2 \leq 0$$

Claims

(i) $\partial_t \mathrm{Ent}(\mu_t) = -I(\mu_t)$

(ii) $|\nabla_- \mathrm{Ent}(\mu_t)|^2 \leq I(\mu_t)$

(iii) $|\dot{\mu}_t|^2 \leq I(\mu_t)$ a.e.t

$$\left(I(\mu_t) := \int_X \frac{|\nabla \rho_t|^2}{\rho_t} dm\text{: Fisher information} \right)$$

$\star$ $|\nabla_d f| = |\nabla f| = |\nabla_- f|$ a.e. for $f \in \mathrm{Lip}(X)$

Claims

(i) $\partial_t \mathrm{Ent}(\mu_t) = -I(\mu_t)$

(ii) $|\nabla_- \mathrm{Ent}(\mu_t)|^2 \le I(\mu_t)$

(iii) $|\dot{\mu}_t|^2 \le I(\mu_t)$ a.e.t

$\left(I(\mu_t) := \int_X \frac{|\nabla \rho_t|^2}{\rho_t} dm\text{: Fisher information} \right)$

⋆ $|\nabla_d f| = |\nabla f| = |\nabla_- f|$ a.e. for $f \in \mathrm{Lip}(X)$

- Integration by parts $\Rightarrow$ (i)
- Directional derivative [Villani '09] & CD(K, ∞) $\Rightarrow$ (ii)

Recall: $|\dot{\mu}_t| := \limsup_{h\downarrow 0} \frac{1}{h} d_2^W(\mu_{t+h}, \mu_t)$

$$\frac{1}{2} d_2^W(\mu_{t+h}, \mu_t)^2$$

$$= \sup_{\varphi\in \mathrm{Lip}(X)} \left[\int_X Q_1\varphi \, d\mu_{t+h} - \int_X \varphi \, d\mu_t\right],$$

(Kantorovich duality)

where $Q_t\varphi(x) := \inf_{y\in X} \left[\varphi(y) + \frac{d(x,y)^2}{2t}\right]$

(Hamilton-Jacobi semigroup)

- [Lott & Villani '07]: $\partial_t Q_t\varphi = -\frac{1}{2}|\nabla Q_t\varphi|^2$ a.e.

$$\int_X Q_1\varphi\, d\mu_{t+h} - \int_X \varphi\, d\mu_t$$

$$= \int_0^1 \partial_r \left(\int_X Q_r\varphi\, d\mu_{t+hr} \right) dr$$

$$\overset{\text{HJ}}{\underset{\text{IbP}}{=}} \int_0^1 dr \int_X d\mu_{t+hr}$$

$$\left(-\frac{1}{2}|\nabla Q_r\varphi|^2 - h \left\langle \nabla Q_r\varphi, \frac{\nabla\rho_{t+hr}}{\rho_{t+hr}} \right\rangle \right)$$

$$\le \frac{h^2}{2} \boxed{\int_0^1 I(\mu_{t+hr}) dr} \qquad \blacksquare$$

4. Applications (under $\mathsf{CD}(K, \infty)$)

L^2-Wasserstein contraction for T_t

$$d_2^W(T_t\mu, T_t\nu) \le \mathrm{e}^{-Kt} d_2^W(\mu, \nu)$$

L^2-Wasserstein contraction for T_t

$$d_2^W(T_t\mu, T_t\nu) \le e^{-Kt} d_2^W(\mu, \nu)$$

$\Downarrow$ [K.'10]

L^2-gradient estimate for $f \in \mathrm{Lip}(X)$

$$|\nabla_d T_t f|^2 \le e^{-2Kt} T_t(|\nabla_d f|^2)$$

L^2-Wasserstein contraction for T_t

$$d_2^W(T_t\mu, T_t\nu) \le e^{-Kt} d_2^W(\mu, \nu)$$

⇓ [K.'10]

L^2-gradient estimate for $f \in \mathrm{Lip}(X)$

$$|\nabla_d T_t f|^2 \le e^{-2Kt} T_t(|\nabla_d f|^2)$$

⇓ $\exists p_t$: conti.

L^2-gradient estimate for $f \in W^{1,2}(X)$

$$|\nabla_d T_t f|^2 \le e^{-2Kt} T_t(|\nabla f|^2)$$

Theorem 2 [Gigli-K.-Ohta]

(i) $T_t f \in \mathrm{Lip}(X)$ for $f \in W^{1,2}(X)$

(ii) For $\forall f$: L^2-eigenfn. of Δ, $f \in \mathrm{Lip}(X)$

(iii) $p_t(x, \cdot) \in \mathrm{Lip}(X)$

- Theorem 2(ii) provides another proof of [Petrunin '03]
- Recently, [Zhang & Zhu] gave another proof of Theorem 2(iii) based on [Petrunin '03]

Theorem 3 [Gigli-K.-Ohta]

For $K_0 \in \mathbb{R}$, the following are equivalent:

(i) $d_2^W(T_t\mu, T_t\nu) \leq e^{-K_0 t} d_2^W(\mu, \nu)$

(ii) For $f \in W^{1,2}(X)$,

$$|\nabla T_t f|^2 \leq e^{-2K_0 t} T_t(|\nabla f|^2) \text{ a.e.}$$

(iii) For $g \in D(\Delta) \cap L^\infty, g \geq 0, \Delta g \in L^\infty$ and $f \in D(\Delta) \cap L^\infty, \Delta f \in W^{1,2}(X)$,

$$\int_X \left(\frac{1}{2} \Delta g \langle \nabla f, \nabla f \rangle - g \langle \nabla f, \nabla \Delta f \rangle \right) dm \geq K_0 \int_X g \langle \nabla f, \nabla f \rangle dm$$

Remarks

(a) Theorem 3 (iii) is nothing but a weak form of Bakry-Émery's Γ_2-condition:

$$\frac{1}{2}\Delta\langle\nabla f,\nabla f\rangle - \langle\nabla f,\nabla\Delta f\rangle \geq K_0\langle\nabla f,\nabla f\rangle$$

(b) $f = T_t\varphi$, $g = T_t\psi$ for some $\varphi, \psi \in L^2$, $\psi \geq 0$ satisfies all requirements on f and g in Theorem 3 (iii).

(c) When X: Riem. mfd, Theorem 3(i)-(iii) are all equivalent to $\mathrm{Ric} \geq K_0$ (or CD(K_0,∞))

[von Renesse & Sturm '05]

Boundedness of Riesz transform on (X, d, m)

[Kawabi & Miyokawa '07]

Theorem 4 [Gigli-K.-Ohta]

Let $2 \le p < \infty$, $q > 1$, and $\alpha > (-K) \vee 0$.

Then

$$\|\nabla(\alpha - \Delta_p)^{-q/2} f\|_{L^p} \le C \|f\|_{L^p}$$