Identification of heat flows

Kazumasa Kuwada

(Ochanomizu University)

非線形偏微分方程式研究会

(2012 年 3 月 5–6 日 早稲田大学)

1. Introduction

Heat equation on $\mathbb{R}^n$:

$$\begin{cases} \dfrac{\partial u}{\partial t} = \Delta u, \\ u(0, \cdot) = f \end{cases}$$

$\star$ $f \in L^2 \Rightarrow u(t, \cdot) \in L^2$

$\star$ $f(x)dx \in \mathcal{P}(\mathbb{R}^n) \Rightarrow u(t, x)dx \in \mathcal{P}(\mathbb{R}^n)$

Two ways to characterize a “heat distribution”

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

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

Two ways to characterize a "heat distribution"

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

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

—— ○ —— ○ ——

On $\mathbb{R}^n$ (or Riem. mfd), based on PDE theory,

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

Two ways to characterize a "heat distribution"

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

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

———— ∘ ———— ∘ ————

On $\mathbb{R}^n$ (or Riem. mfd), based on PDE theory,

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

$$\Rightarrow \boxed{(1) = (2)}$$

Q. What happens on more singular spaces?

Q. What happens on more singular spaces?

- No (second order) differentiable structure

◦ Both (1) and (2) require only "1st order calc."

$\Updownarrow$ metric space structure

Q. What happens on more singular spaces?

- No (second order) differentiable structure

◦ Both (1) and (2) require only "1st order calc."
  $\Uparrow$ metric space structure

Ans.

(1) & (2) coincide on metric measure sp. (X, d, m) where 1st order calc. works well

Known results

- $\mathbb{R}^n$ [Jordan, Kinderlehrer & Otto '98]
- cpl. Riem. mfd [Erbar '10]
- Finsler mfd [Ohta & Sturm '09] (nonlinear)
- Wiener sp. [Fang, Shao & Sturm '09] (∞-dim)
- Heisenberg gr. [Juillet] (sub-elliptic)
- finite set [Maas '11] (discrete Markov chain)

——— ∘ ——— ∘ ———

- cpt. Alexandrov sp. [Gigli, K. & Ohta] (singular)
- geodesic mm-sp. [Ambrosio, Gigli & Savaré]

Identification of (1) and (2)

$$\Downarrow$$

Properties from (1) + Properties from (2)

(under a lower curvature bound)

Identification of (1) and (2)

$\Downarrow$

Properties from (1) + Properties from (2)

(under a lower curvature bound)

$\Downarrow$

Theorem [Gigli, K. & Ohta]

On a cpt. Alexandrov sp. without boundary,
the heat kernel $p_t(x, \cdot)$ is Lipschitz continuous

(Approach only from (1) $\Rightarrow$ Hölder continuity)

2. Formulation of gradient flows on Riem. mfd

X: cpt. Riemannian mfd, $\partial X = \emptyset$

d: Riem. distance function

m: Riem. volume measure

$\star$ Suppose $\mathrm{Ric} \geq K$ for some $K \in \mathbb{R}$

2.1. Dirichlet energy and its gradient flow

$(\mathcal{E}, H^{1,2}(X))$: Dirichlet energy on $L^2(X, m)$

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

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

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

$$\Downarrow$$

$T_t f$ solves the heat equation

(in the sense of evolution equation)

†

$$\lim_{\delta\to 0}\frac{1}{2\delta}\left(\mathcal{E}(f+\delta g)-\mathcal{E}(f)\right) = -\int_X \Delta f\, g\, dm$$

$$\Rightarrow \boxed{\nabla\left(\frac{\mathcal{E}}{2}\right)(f) = -\Delta f}$$

$\Rightarrow$ a gradient curve

$$\partial_t u = -\nabla\left(\frac{\mathcal{E}}{2}\right)(u), \quad u(0,\cdot) = f$$

on $L^2(X,m)$ solves the heat eq.

$\Rightarrow$ $T_t f$: gradient curve of $\mathcal{E}/2$

† $\mathrm{Lip}(X) \subset H^{1,2}(X)$ dense

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

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

2.2. Formal Riemannian structure on $(\mathcal{P}_2(X), W_2)$

L^2-Wasserstein distance

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

$\Pi(\mu_0, \mu_1) \subset \mathcal{P}(X \times X)$: coupling of μ_0 & μ_1

$$\Pi(\mu_0, \mu_1) := \left\{ \pi \;\middle|\; \begin{array}{l} \pi(A \times X) = \mu_0(A) \\ \pi(X \times A) = \mu_1(A) \end{array} \right\}$$

$$W_2(\mu_0, \mu_1) := \inf_{\pi \in \Pi(\mu,\nu)} \|d\|_{L^2(\pi)}$$

L^2-Wasserstein distance

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

$\Pi(\mu_0, \mu_1) \subset \mathcal{P}(X \times X)$: coupling of μ_0 & μ_1

$$\Pi(\mu_0, \mu_1) := \left\{ \pi \,\middle|\, \begin{array}{l} \pi(A \times X) = \mu_0(A) \\ \pi(X \times A) = \mu_1(A) \end{array} \right\}$$

$$W_2(\mu_0, \mu_1) := \inf_{\pi \in \Pi(\mu, \nu)} \|d\|_{L^2(\pi)}$$

- $\mathcal{P}_2(X) := \{\mu \mid W_2(\delta_x, \mu) < \infty\} = \mathcal{P}(X)$

L^2-Wasserstein distance

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

$\Pi(\mu_0, \mu_1) \subset \mathcal{P}(X \times X)$: coupling of μ_0 & μ_1

$$\Pi(\mu_0, \mu_1) := \left\{ \pi \;\middle|\; \begin{array}{l} \pi(A \times X) = \mu_0(A) \\ \pi(X \times A) = \mu_1(A) \end{array} \right\}$$

$$\boxed{W_2(\mu_0, \mu_1) := \inf_{\pi \in \Pi(\mu,\nu)} \|d\|_{L^2(\pi)}}$$

- $\mathcal{P}_2(X) := \{\mu \mid W_2(\delta_x, \mu) < \infty\} = \mathcal{P}(X)$

$\star$ $(\mathcal{P}_2(X), W_2)$: 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.

- $\mathrm{e}_0^{\#}\Pi = \mu_0$, $\mathrm{e}_1^{\#}\Pi = \mu_1$
- $W_2(\mathrm{e}_t^{\#}\Pi, \mathrm{e}_s^{\#}\Pi) = |t - s| W_2(\mu_0, \mu_1)$

([Lott & Villani '09] when X: geodesic sp.)

More detailed description of W_2 on Riem. mfd

$\mu_0, \mu_1 \in \mathcal{P}_2(X)$, $\mu_0 \ll m$

$(\mu_t)_{t\in[0,1]}$: W_2-min. geod.

$\Rightarrow {}^{\exists}\varphi : X \to \mathbb{R}$: convex s.t.

- $\exp(t\nabla\varphi)^{\#}\mu_0 = \mu_t$,
- $W_2(\mu_0, \mu_t)^2 = t^2 \int_X |\nabla\varphi|^2 d\mu_0$

[Brenier '91, McCann '95]

Formal Riem. structure on $\mathcal{P}_2(X)$ [Otto '01]

- Tangent space at $\mu \in \mathcal{P}_2(X)$:

$$T_\mu \mathcal{P}_2(X) := \overline{\{\nabla \varphi \mid \varphi \in C^\infty(X)\}}^{L^2(\mu)}$$

- Riem. metric on $T_\mu \mathcal{P}_2(X)$:

$$\sigma(\nabla\varphi, \nabla\psi)(\mu) := \int_{\mathbb{R}^m} \langle \nabla\varphi, \nabla\psi \rangle d\mu$$

"Regular" curve in $\mathcal{P}_2(X)$

$\varphi_t \in C_0^\infty(X)$, Φ_t: grad. flow of $-\varphi_t$ on X

$\mu_t := \Phi_t^\# \mu \;(\Rightarrow \nabla \varphi_t \in T_{\mu_t}\mathcal{P}_2(\mathbb{R}^m))$

$$\Rightarrow \frac{d}{dt}\int_X f\, d\mu_t = \frac{d}{dt}\int_X f \circ \Phi_t\, d\mu$$

$$= \int_X \langle (\nabla f) \circ \Phi_t, \partial_t \Phi_t \rangle d\mu$$

$$= \int_X \langle \nabla f, \nabla \varphi_t \rangle d\mu_t$$

$$\Rightarrow \frac{d}{dt}\mu_t + \mathrm{div}_{\mu_t}(\nabla \varphi_t)\mu_t = 0$$

2.3. Gradient flow of relative entropy on $(\mathcal{P}_2(X), W_2)$

W_2-gradient of Ent

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

$$\frac{d}{dt}\operatorname{Ent}(\mu_t)|_{t=0} = \frac{d}{dt}\int_X \log\rho_t \, d\mu_t\Big|_{t=0}$$

$$= \int_X \partial_t\rho_0 \, dm + \int_X \langle \frac{\nabla\rho_0}{\rho_0}, \nabla\varphi_0\rangle d\mu$$

$$= \sigma(\frac{\nabla\rho_0}{\rho_0}, \nabla\varphi_0)(\mu)$$

$$\Rightarrow \boxed{\nabla\operatorname{Ent}(\mu) = \frac{\nabla\rho}{\rho}}$$

W_2-Gradient flow of Ent

- $\dfrac{d}{dt}\mu_t + \mathrm{div}_{\mu_t}(\nabla\varphi_t)\mu_t = 0$
- $\dfrac{d}{dt}\mu_t = -\nabla\,\mathrm{Ent}(\mu_t)$ iff $\nabla\varphi_t = -\dfrac{\nabla\rho_t}{\rho_t}$

$\Rightarrow$ When μ_t: grad. curve of $-\,\mathrm{Ent}$,

$$\frac{d}{dt}\int_X f\,d\mu_t = -\int_X \langle\nabla f, \nabla\rho_t\rangle dm$$

$$= \int_X \Delta f\,d\mu_t$$

$\therefore$ μ_t solves the heat equation (weakly)

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

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

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

$$|\dot{\mu}_r| := \limsup_{h\downarrow 0} \frac{1}{h} W_2(\mu_{r+h}, \mu_r)$$

$$|\nabla_- \mathrm{Ent}(\mu)| := \limsup_{\nu\to\mu} \frac{[\mathrm{Ent}(\mu) - \mathrm{Ent}(\nu)]_+}{W_2(\mu,\nu)}$$

Heuristics:

Why does this definition work?

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

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

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

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

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

The condition CD(K,∞)

For $\forall(\nu_t)_{t\in[0,1]}$: W_2-min. geod.,

$$\mathrm{Ent}(\nu_\lambda) \le (1-\lambda)\,\mathrm{Ent}(\nu_0) + \lambda\,\mathrm{Ent}(\nu_1) - \frac{K}{2}\lambda(1-\lambda)W_2(\nu_0,\nu_1)^2$$

(K-convexity of Ent w.r.t. W_2)

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

For $\forall (\nu_t)_{t\in[0,1]}$: W_2-min. geod.,

$$\mathrm{Ent}(\nu_\lambda) \le (1-\lambda)\,\mathrm{Ent}(\nu_0) + \lambda\,\mathrm{Ent}(\nu_1) - \frac{K}{2}\lambda(1-\lambda)W_2(\nu_0,\nu_1)^2$$

(K-convexity of Ent w.r.t. W_2)

- When X: Riem. mfd,

$$\mathrm{CD}(K,\infty) \Leftrightarrow \mathrm{Ric} \ge K$$

[von Renesse & Sturm '05]

Existence and uniqueness of gradient flow

Under $\mathsf{CD}(K,\infty)$,

$\exists!$ grad. flow of Ent on $(\mathcal{P}_2(X), W_2)$

starting from $\forall\ \mu \in \mathcal{P}_2(X)$ with $\mathrm{Ent}(\mu) < \infty$

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

(In fact, $\mathsf{CD}(K,\infty)$ is not necessary

[Ambrosio, Gigli & Savaré])

3. Identification

We are still in the framework of §2

Theorem 1

For any $f \in L^2$ with $fm \in \mathcal{P}_2(X)$,
$(T_t f)m \in \mathcal{P}_2(X)$ is a gradient flow of Ent

Notation:

$$(T_t f)m =: \mu_t, \quad T_t f =: \rho_t$$

Goal

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

- "$\geq$" is always true
- Suffices to show:

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

for a.e.t

Claims

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

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

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

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

Claims

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

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

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

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

Recall: $|\nabla_d f| = |\nabla f|$ a.e. for $f \in \mathrm{Lip}(X)$

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

Kantorovich duality

$$\frac{1}{2}W_2(\mu,\nu)^2 = \sup_{g,f}\left[\int_X g\,d\mu + \int_X f\,d\nu\right]$$
$$= \sup_{f}\left[\int_X \hat{f}\,d\mu + \int_X f\,d\nu\right],$$

where $f, g \in C_b(X)$,

$$g(x) + f(y) \leq \frac{1}{2}d(x,y)^2,$$

$$\hat{f}(x) := \inf_{y\in X}\left[\frac{1}{2}d(x,y)^2 - f(y)\right]$$

Constraint: $g(x) + f(y) \leq \frac{1}{2} d(x,y)^2$

$\Rightarrow$ For $\pi \in \Pi(\mu, \nu)$,

$$\int_X g \, d\mu + \int_X f \, d\nu$$

$$= \int_X (g(x) + f(y)) \pi(dx dy) \leq \frac{1}{2} \|d\|^2_{L^2(\pi)}$$

$$\Rightarrow \frac{1}{2} W_2(\mu, \nu)^2 \geq \sup_{g,f} \left[\int_X g \, d\mu + \int_X f \, d\nu \right]$$

Recall: $|\dot\mu_t| := \limsup_{h\downarrow 0} \frac{1}{h} W_2(\mu_{t+h}, \mu_t)$

$$\frac{1}{2} W_2(\mu_{t+h}, \mu_t)^2$$

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

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

(Hamilton-Jacobi/Hopf-Lax semigroup)

⋆ $\partial_s Q_s\varphi = -\frac{|\nabla_d Q_s\varphi|^2}{2}$ a.e. [Lott & Villani '07]

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

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

$$\overset{\text{HJ, IbP}}{=} \int_0^1 dr \int_X d\mu_{t+hr} \left(-\frac{1}{2}|\nabla Q_r\varphi|^2 - h \left\langle \nabla Q_r\varphi, \frac{\nabla \rho_{t+hr}}{\rho_{t+hr}} \right\rangle \right)$$

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

4. Extensions

(A) cpt. Alexandrov sp. [Gigli, K. & Ohta]

(B) geodesic mm-sp. (X, d, m)

[Ambrosio, Gigli & Savaré]

- ⋆ d may NOT be compatible with the (Polish) topology of X (but related)
- ⋆ $\int_X \mathrm{e}^{-\kappa d^2(o,x)} m(dx) < \infty$

  for some $o \in X$, $\kappa > 0$

  (when d: dist. being compatible with the top.)
- ⋆ CD (K, ∞) is NOT assumed

Dirichlet energy f'nal

(A) – a.e. differentiable structure

[Otsu & Shioya '94]

– bilinear Dirichlet energy (Dirichlet form)

[Kuwae, Machigashira & Shioya '01]

(B) – (modified) Cheeger-type energy via upper gradient (possibly nonlinear)

Dirichlet energy f'nal

(A) – a.e. differentiable structure

[Otsu & Shioya '94]

– bilinear Dirichlet energy (Dirichlet form)

[Kuwae, Machigashira & Shioya '01]

(B) – (modified) Cheeger-type energy via upper gradient (possibly nonlinear)

† Relation between energy density & loc. Lip const.

† $\mathrm{Lip}(X)$: dense in the domain of the D-energy

† Integration by parts

Identification under the lack of CD (K, ∞)

- Work with $I(\mu)$ instead of $|\nabla_- \mathrm{Ent}|(\mu)$ in the def. of GF of Ent
- For a GF of Ent,

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

- The same with "$\geq$" for a GF of Dirichlet energy
- Uniqueness argument in [Gigli '10]

5. Applications (under $\mathrm{CD}(K,\infty)$)

⋆ CD (K, ∞) for some $K \in \mathbb{R}$ on an Alex. sp.

[Petrunin '11]

⋆ CD (K, ∞) for some $K \in \mathbb{R}$ on an Alex. sp.

[Petrunin '11]

⋆ CD (K, ∞) ($\Leftrightarrow$ "$\nabla^2 \operatorname{Ent} \geq K$")

$\Rightarrow$ For μ_t & ν_t: GFs of Ent,

$$W_2(\mu_t, \nu_t) \leq e^{-Kt} W_2(\mu_0, \nu_0)$$

if the Dirichlet energy is bilinear

[Ambrosio, Gigli & Savaré]

(On Alex. sp., [Ohta '09, Savaré '07])

L^2-Wasserstein contraction for T_t

$$W_2(T_t\mu, T_t\nu) \leq \mathrm{e}^{-Kt} W_2(\mu, \nu)$$

L^2-Wasserstein contraction for T_t

$$W_2(T_t\mu, T_t\nu) \le \mathrm{e}^{-Kt} W_2(\mu, \nu)$$

$\Downarrow$ [cf. K.'10]

Bakry-Émery's L^2-gradient estimate

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

Theorem 2 [Gigli, K. & Ohta]

On cpt. Alex. sp.,

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

(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 [Gigli, K. & Ohta]

On cpt. Alex. sp.,

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

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

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

- Extension by [Ambrosio, Gigli & Savaré]
- Theorem 2(ii): another proof of [Petrunin '03]
- [Zhang & Zhu]: another proof of Theorem 2(iii) based on [Petrunin '03]