Duality on gradient estimates and Wasserstein controls

Kazumasa Kuwada

(Ochanomizu University
Universität Bonn)

§1 Introduction

Equivalent conditions for a lower Ricci curvature bound

(von Renesse & Sturm '05, etc...)

X: complete Riemannian manifold

P_t: heat semigroup associated with Δ

(i) $\mathrm{Ric} \geq k$

(ii) $d_p^W(P_t^*\mu, P_t^*\nu) \leq \mathrm{e}^{-kt} d_p^W(\mu, \nu)$
for some $p \in [1, \infty]$

(iii) $|\nabla P_t f|(x) \leq \mathrm{e}^{-kt} P_t(|\nabla f|^q)(x)^{1/q}$
for some $q \in [1, \infty]$

Our goal:

Generalization of (ii) $\Leftrightarrow$ (iii)

for p, q with $\frac{1}{p} + \frac{1}{q} = 1.$

(ii) L^p-Wasserstein control

$$d_p^W(P_t^* \mu, P_t^* \nu) \le \mathrm{e}^{-kt} d_p^W(\mu, \nu)$$

(iii) L^q-gradient estimate

$$|\nabla P_t f|(x) \le \mathrm{e}^{-kt} P_t(|\nabla f|^q)(x)^{1/q}$$

§2 Framework and main result

(X, d): Polish metric space.

- $(P_x)_{x \in X} \subset \mathcal{P}(X)$: Markov kernel.

$$Pf(x) := \int_X f \, dP_x,$$

$$P^* \mu(A) := \int_X P_x(A) \mu(dx).$$

(e.g. $P = P_t$: heat semigroup)

- $\tilde{d}$: continuous distance function on X.
(e.g. $\tilde{d} = \mathrm{e}^{-kt} d$)

$\Pi(\mu,\nu)$: set of couplings for $\mu,\nu \in \mathcal{P}(X)$, i.e.

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

L^p-Wasserstein (pseudo) distance

For $p \in [1,\infty]$,

$$d_p^W(\mu,\nu) := \inf_{\pi \in \Pi(\mu,\nu)} \|d\|_{L^p(\pi)} \in [0,\infty].$$

Gradient

$$|\nabla_d f|(x) := \lim_{r \downarrow 0} \sup_{y \in B_r(x)} \left| \frac{f(x) - f(y)}{d(x, y)} \right|$$

$$\|\nabla_d f\|_\infty := \sup_{x \in X} |\nabla_d f|(x)$$

Gradient

$$|\nabla_d f|(x) := \lim_{r \downarrow 0} \sup_{y \in B_r(x)} \left| \frac{f(x) - f(y)}{d(x,y)} \right|$$

$$\|\nabla_d f\|_\infty := \sup_{x \in X} |\nabla_d f|(x)$$

f: Lipschitz $\Rightarrow$ $|\nabla_d f|$ is an upper gradient of f

i.e. $$f(y) - f(x) \le \int_a^b |\nabla_d f|(\gamma(s))\,ds,$$

$\forall \gamma \; : \; [a,b] \to X$ 1-Lipschitz curve from x to y

L^p-Wasserstein control

$$d_p^W(P^*\mu, P^*\nu) \le \tilde{d}_p^W(\mu, \nu) \qquad (W_p)$$

for $p \in [1, \infty]$ and $\mu, \nu \in \mathcal{P}(X)$

L^q-gradient estimate

$$|\nabla_{\tilde{d}} Pf|(x) \le P(|\nabla_d f|^q)(x)^{1/q} \qquad (G_q)$$

for $q \in [1, \infty)$ and $f \in C_b^{\mathrm{Lip}}(X)$,

$$\|\nabla_{\tilde{d}} Pf\|_\infty \le \|\nabla_d f\|_\infty \qquad (G_\infty)$$

for $q = \infty$

v: Radon measure on X with $\mathrm{supp}(v) = X$.

Assumption 1 (X, d): proper length space

Assumption 2 (X, d, v) supports

- local (uniform) volume doubling condition
- $(1, \rho)$-local Poincaré inequality $(\exists \rho \geq 1)$

Assumption 3 $\tilde{d}$: geodesic distance

Assumption 4 $P_x \ll v$, $x \mapsto \dfrac{dP_x}{dv}(y)$: continuous

v: Radon measure on X with $\mathrm{supp}(v) = X$.

Assumption 1 (X, d): proper length space

Assumption 2 (X, d, v) supports

- local (uniform) volume doubling condition
- $(1, \rho)$-local Poincaré inequality $(\exists \rho \geq 1)$

Assumption 1,2 enable us to employ general theory of Hamilton-Jacobi semigroups

Local uniform volume doubling condition

$\exists D > 0,\ \exists R_1 > 0$ s.t. $\forall x \in X$, $\forall r < R_1$

$$v(B_{2r}(x)) \leq D v(B_r(x)).$$

$(1,\rho)$-local Poincaré inequality

$\forall R_2 > 0,\ \exists \lambda \geq 1,\ \exists C_P > 0$ s.t. $\forall r < R_2$,

$$⨍_{B_r(x)} |f - f_{x,r}| dv \leq C_P r \left(⨍_{B_{\lambda r}(x)} g^\rho \, dv \right)^{1/\rho}$$

for $\forall f$ and $\forall g$: upper gradient of f, where

$$f_{x,r} := \frac{1}{v(B_r(x))} \int_{B_r(x)} f \, dv =: ⨍_{B_r(x)} f \, dv.$$

Theorem (K. '09)

For $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$,

(i) $(W_p) \Rightarrow (G_q)$

(ii) Under Assumption 1-4, $(G_q) \Rightarrow (W_p)$

Remarks

- For $p' > p$, $\begin{cases} (G_p) \Rightarrow (G_{p'}) \\ (W_{p'}) \Rightarrow (W_p) \end{cases}$
  (without Assumption 1-4)

- $(G_\infty) \Leftrightarrow (W_1)$ is well known.
  $\left(\begin{array}{l} \text{via Kantorovich-Rubinstein formula;} \\ \text{without Assumption 1-4} \end{array} \right)$

- $(W_\infty) \Rightarrow (G_1)$ is essentially well known.

§3 Examples and applications

How do we obtain (W_p)?

(A) Two known derivations of (W_p)

(1) Coupling by parallel transport of B.m.'s

X: cpl. Riem. mfd., $\mathbf{Ric} \geq k$

$\exists (B_t, \tilde{B}_t)$: coupling of B.m's s.t.

$$d(B_t, \tilde{B}_t) \leq \mathrm{e}^{-kt/2} d(B_0, \tilde{B}_0) \quad \mathbb{P}\text{-a.s.}$$

i.e. $\boxed{\mathbf{Ric} \geq k \Rightarrow (W_\infty)}$

Extension:

Backward (super) Ricci flow $\partial_t g(t) \leq \mathrm{Ric}_{g(t)}$

For $g(t)$-B.m.'s ($\leftrightarrow$ generator $\partial_t + \frac{1}{2}\Delta_{g(\cdot)}$),

$$d_{g(t)}(B_t, \tilde{B}_t) \leq d_{g(s)}(B_s, \tilde{B}_s) \quad \mathbb{P}\text{-a.s.}$$

- McCann & Topping '08: (W_2), X: cpt
- Arnaudon, Coulibaly & Thalmaier '09: (W_∞) (cf. K. & Philipowski '09 for non-explosion)

(2) Gradient flow formulation of the heat flow μ_t

X: cpl. Riem. mfd.

"$\partial_t \mu_t = -\nabla E(\mu_t)$" ($E$: relative entropy)

- $\mathrm{Ric} \geq k \Leftrightarrow$ "$\mathrm{Hess}\, E \geq k$",
- $\mathrm{Hess}\, E \geq k \Rightarrow (W_2)$ for μ_t $(= P_t^* \mu)$

(Heuristically, differential geometry on $\mathcal{P}(X)$)

$\rightsquigarrow$ Extension to singular spaces

(e.g. Alexandrov spaces)

Remark

"(a lower Ricci bound) $\Rightarrow (W_p)$" in the literature.
No direct way "$(G_q) \Rightarrow (W_p)$" was known.

E.g. in von Renesse & Sturm '05,

$$
\begin{array}{ccccccc}
\mathrm{Ric} \geq k & & & & & & \\
\Downarrow \ \text{coupling method} & & & & & & \\
(W_\infty) & \Rightarrow & (W_p) & \Rightarrow & (W_1) & & \\
\Downarrow & & & & \Downarrow & & \\
(G_1) & \Rightarrow & (G_q) & \Rightarrow & (G_\infty) & \Rightarrow & \mathrm{Ric} \geq k \\
 & & & & & \text{Bochner} &
\end{array}
$$

(B) Hörmander-type operators on a Lie group

X: Lie group

$\{X_i\}_{i=1}^n$: left-invariant, lin. indep. vector fields satisfying the Hörmander condition

$$P_t := \mathrm{e}^{tA}, \quad A := \sum_{i=1}^n X_i^2,$$

$$|\Gamma f| := \frac{1}{2}\bigl(A(f^2) - 2fAf\bigr) = \sum_{i=1}^n |X_i f|^2$$

<u>L^q-gradient estimate</u>

$$|\Gamma P_t f|(x) \le K_q(t) P_t(|\Gamma f|^{q/2})(x)^{2/q} \qquad (G_q^*)$$

Known results

- 3-dim. Heisenberg group, $K_q(t) \equiv K_q > 1$
  - $q > 1$: Driver & Melcher '05
  - $q = 1$: H.-Q. Li '06 / Bakry, Baudoin, Bonnefont & Chafaï '08
- X: general, $q > 1$: Melcher '08 ($K_q(t) \equiv K_q$ if X: nilpotent)
- X: group of type H, $q = 1$, $K_q(t) \equiv K_q$: Eldredge '10
- $X = SU(2)$, $q > 1$, $K_q(t) = K_q \mathrm{e}^{-t}$: Baudoin & Bonnefont '09

Carnot-Caratheodory distance

For $V \in T_x X$,

$$|V| = \begin{cases} \left(\sum_{i=1}^{n} {a_i}^2 \right)^{1/2} & \text{if } V = \sum_{i=1}^{n} a_i X_i(x), \\ \infty & \text{otherwise.} \end{cases}$$

$$d(x, y) := \inf \left\{ \int_0^1 |\dot{\gamma}_s| ds \;\middle|\; \begin{array}{l} \gamma_0 = x, \\ \gamma_1 = y \end{array} \right\}$$

v: right-Haar measure, $P = P_t$.

Proposition

(i) $(X, d, v; P)$ satisfies Assumption 1-4

(ii) $(G_q^*) \Rightarrow (G_q)$

Corollary

$(G_q^*) \Rightarrow (W_p)$ for $q \in [1, \infty]$.

Examples

3-dim. Heisenberg group

$X = \mathbb{R}^3$, v: Lebesgue

$$(x, y, z) \cdot (x', y', z') = \left(x + x', y + y', z + z' + \frac{1}{2}(xy' - yx')\right)$$

$$X_1 = \partial_x - \frac{y}{2}\partial_z , \quad X_2 = \partial_y + \frac{x}{2}\partial_z$$

Associated diffusion (B_t^1, B_t^2, B_t^3) from (x, y, z):

$$B_t^1 = W_t^1, \qquad B_t^2 = W_t^2,$$

$$B_t^3 = z + \frac{1}{2}\int_0^t W_t^1 dW_t^2 - W_t^2 dW_t^1,$$

where (W_t^1, W_t^2): 2-dim. BM from (x, y)

(W_∞): For each $t > 0$,

$\exists$ a coupling $(\boldsymbol{B}_t, \tilde{\boldsymbol{B}}_t)$ of (B_t^1, B_t^2, B_t^3) s.t.

$$d(\boldsymbol{B}_t, \tilde{\boldsymbol{B}}_t) \le K_1 d(\boldsymbol{B}_0, \tilde{\boldsymbol{B}}_0) \quad \mathbb{P}\text{-a.s.}$$

Definition

X: a group of type H iff, for $\mathcal{X}$: Lie alg. associated with X with a scalar product $\langle \cdot, \cdot \rangle$,

- $\mathcal{X} = \mathcal{V} \oplus \mathcal{Z}$ with $[\mathcal{V}, \mathcal{V}] = \mathcal{Z}$, $[\mathcal{V}, \mathcal{Z}] = [\mathcal{Z}, \mathcal{Z}] = 0$.
- $J : \mathcal{Z} \to \mathrm{End}\, \mathcal{V}$ given by

$$\langle J(Z)V_1, V_2 \rangle := \langle Z, [V_1, V_2] \rangle$$

satisfies $J(Z)^2 = -\|Z\| \mathrm{Id}$.

$\{X_i\}_{i=1}^n$ will be an ONB of $\mathcal{V}$.

Remarks

- (Heisenberg) $\subset$ (type H)

Remarks

- (Heisenberg) $\subset$ (type H)
- (type H) $\subset$ (stratified, step 2 nilpotent)
- (type H) $\cap$ (free step 2 nilpotent)

  $= \{$3-dim. Heisenberg$\}$
- For type H,

  possible dimension is completely determined

§4 Sketch of the proof of $(G_q) \Rightarrow (W_p)$

Recall:

$$d_p^W(P^*\mu, P^*\nu) \leq \tilde{d}_p^W(\mu, \nu) \qquad (W_p)$$

$$|\nabla_{\tilde{d}} Pf|(x) \leq P(|\nabla_d f|^q)(x)^{1/q} \qquad (G_q)$$

Recall:

$$d_p^W(P^*\mu, P^*\nu) \le \tilde{d}_p^W(\mu,\nu) \qquad (W_p)$$

$$|\nabla_{\tilde{d}} Pf|(x) \le P(|\nabla_d f|^q)(x)^{1/q} \qquad (G_q)$$

- $p = 1$ $(q = \infty)$: well-known
- (W_p) for $\forall p < \infty \Rightarrow (W_\infty)$

  $\rightsquigarrow$ We may assume $p \in (1, \infty)$

Recall:

$$d_p^W(P^*\mu, P^*\nu) \le \tilde{d}_p^W(\mu,\nu) \qquad (W_p)$$

$$|\nabla_{\tilde{d}} Pf|(x) \le P(|\nabla_d f|^q)(x)^{1/q} \qquad (G_q)$$

- $p = 1$ $(q = \infty)$: well-known
- (W_p) for $\forall p < \infty \Rightarrow (W_\infty)$

  $\rightsquigarrow$ We may assume $p \in (1, \infty)$

○ (W_p) for $\mu = \delta_x, \nu = \delta_y \Rightarrow (W_p)$

  $\rightsquigarrow$ We show $\dfrac{d_p^W(P_x, P_y)^p}{p} \le \dfrac{\tilde{d}(x,y)^p}{p}$

General theory of the Hamilton-Jacobi semigroup

(Lott & Villani '07, Balogh, Engoulatov, Hunziker & Maasalo '09)

$$Q_t f(x) := \inf_{y \in X} \left[f(y) + t \cdot \frac{1}{p} \left(\frac{d(x,y)}{t} \right)^p \right]$$

- Under Assumption 1, $Q_\cdot f \in C_b^{\mathrm{Lip}}([0,\infty) \times X)$ if $f \in C_b^{\mathrm{Lip}}(X)$
- Under Assumption 1-2, for $\forall t > 0$, υ-a.e.

$$\boxed{\partial_t Q_t f = -\frac{1}{q} |\nabla_d Q_t f|^q}$$

(Note: $q^{-1} u^q = \sup_{s \geq 0} \left(us - p^{-1} s^p \right)$)

Kantorovich duality

$$d_p^W(\mu,\nu)^p = \sup_{f\in C_b^{\mathrm{Lip}}} \left[\int_X f^*\, d\mu - \int_X f\, d\nu\right],$$

$$f^*(x) := \inf_{y\in X} [\, f(y) + d(x,y)^p \,]$$
$$= p\, Q_1(p^{-1} f)(x)$$

$$\Downarrow$$

$$\frac{d_p^W(\mu,\nu)^p}{p} = \sup_f \left[\int_X Q_1 f\, d\mu - \int_X f\, d\nu\right]$$

$$\begin{cases} \tilde{\gamma} : [0,1] \to X \quad \tilde{d}\text{-minimal geodesic}, \\ \tilde{\gamma}_0 = y,\ \tilde{\gamma}_1 = x, \\ \tilde{d}(\tilde{\gamma}_s, \tilde{\gamma}_t) = |t-s|\tilde{d}(x,y) \end{cases}$$

(Assumption 3)

—— ○ —— ○ ——

$$\frac{d_p^W(P_x, P_y)^p}{p} = \sup_f \left[PQ_1 f(x) - Pf(y)\right]$$

interpolation

$$= \sup_f \left[\int_0^1 \partial_t (PQ_t f(\tilde{\gamma}_t)) dt\right]$$

$$
\begin{aligned}
&\partial_t (PQ_t f(\tilde{\gamma}_t)) \\
&\quad \left(\text{“}{=}\text{”}\ P(\partial_t Q_t f)(\tilde{\gamma}_t) + \langle \nabla PQ_t f(\tilde{\gamma}_t), \dot{\tilde{\gamma}}_t \rangle \right) \\
&\underset{\text{up. grad.}}{\overset{\text{HJ eq.}}{\boxed{\le}}} -\frac{1}{q} P(|\nabla_d Q_t f|^q)(\tilde{\gamma}_t) \\
&\qquad\qquad + \tilde{d}(x,y) \left| \nabla_{\tilde{d}} PQ_t f \right| (\tilde{\gamma}_t) \\
&\overset{(G_q)}{\boxed{\le}} \tilde{d}(x,y)\sigma - \frac{1}{q}\sigma^q \le \frac{\tilde{d}(x,y)^p}{p}.
\end{aligned}
$$

$$
\left(\sigma := P(|\nabla_d Q_t f|^q)(\tilde{\gamma}_t)^{1/q} \right) \qquad \blacksquare
$$

§5 Duality under different assumptions

$\boldsymbol{X}$: Polish space

$\boldsymbol{d}, \tilde{\boldsymbol{d}}$: lower semicontinuous pseudo-distances

$\boldsymbol{P}$: Markov kernel, $\boldsymbol{P(C_b(X)) \subset C_b(X)}$

Assumption 5

For $\forall x, y \in X$ with $d(x, y) < \infty$,

$\exists\gamma$: minimal geodesic from x to y.

The same is true for $\tilde{d}$

Assumption 6

$Q_t f$ is measurable for $\forall f \in C_b(X)$

Assumption 7

$$\lim_{r \downarrow 0} \sup_{y;\ \tilde{d}(x,y) \leq r} PQ_t f(y) \leq PQ_t f(x)$$

“Assumption 8”

The following holds locally:

- $\exists D \geq 0$ s.t.
  for $\forall\gamma$: d-min. geod., $\forall\lambda \in [0, 1]$,

$$d(x, \gamma(\lambda))^2 \geq (1-\lambda)d(x, \gamma(0))^2 + \lambda d(x, \gamma(1))^2 - D\lambda(1-\lambda)d(\gamma(0), \gamma(1))^2 \qquad (S(D))$$

- $\forall y, x \in X$, with $d(x, y)$: small,
  y is on a min. geod. from x of given length

Assumption 8

For $\forall \boldsymbol{K} \subset \boldsymbol{X}$ cpt., $\exists \boldsymbol{D_K} \geq 0, \exists \boldsymbol{\eta_K} > 0$ s.t.

(i) for $\forall \gamma$: d-min. geod. with $\gamma(0) \in \boldsymbol{K}$, $\boldsymbol{d}(\gamma(0), \gamma(1)) < \eta_K$, $\boldsymbol{d}(\gamma(0), \boldsymbol{x}) < \eta_K$ and $\boldsymbol{\lambda} \in (0, 1)$, $(\boldsymbol{S}(\boldsymbol{D_K}))$ holds.

(ii) For $\forall \boldsymbol{x} \in \boldsymbol{K}$, $\forall \boldsymbol{y} \in \boldsymbol{X}$ with $\boldsymbol{d}(\boldsymbol{x}, \boldsymbol{y}) < \eta_K$, $\exists \gamma$: min. geod. with $\boldsymbol{d}(\boldsymbol{x}, \gamma(1)) = \eta_K$ and $\boxed{\gamma(\boldsymbol{\lambda}) = \boldsymbol{y}}$ for some $\boldsymbol{\lambda} \in (0, 1)$.

Examples

- X: cpl. Riem. mfd., d: Riem. distance
  - Ass. 8 (i) $\Leftarrow$ local lower sect. curv. bound
  - Ass. 8 (ii) $\Leftarrow$ local positivity of inj. radius
- X: Wiener space, d: Cameron-Martin norm
  - Ass. 8 (i) holds with $D_K = 1$ ("$=$" holds!)
  - Ass. 8 (ii) is obvious

Subgradient

$$|\nabla_d^- f|(x) := \lim_{r\downarrow 0} \sup_{y\in B_r(x)} \left[\frac{f(x)-f(y)}{d(x,y)}\right]_+$$

$$|\nabla_{\tilde{d}}^- Pf|(x) \le P(|\nabla_d^- f|^q)(x)^{1/q} \qquad (G_q^-)$$

Theorem (K. '10)

For $p, q \in [1,\infty]$ with $\frac{1}{p}+\frac{1}{q}=1$,

(i) $(W_p) \Rightarrow (G_q^-)$

(ii) Under Assumption 5-8, $(G_q^-) \Rightarrow (W_p)$

Difficulty: Leibniz rule for $PQ_t f(\tilde{\gamma}(t))$

$$\left(\begin{array}{r} \text{Neither } s \mapsto PQ_t f(\tilde{\gamma}(s)) \\ \text{nor } s \mapsto PQ_s f(\tilde{\gamma}(t)) \text{ is of class } C^1 \end{array} \right)$$

Key of the proof

Assumption 8

$\Rightarrow$ sharp (local) uniform upper bound of

$$\frac{Q_{t+s} f - Q_s f}{t} \text{ for small } t$$

§6 Questions

(i) When (weak) $\Rightarrow$ (strong) ?
i.e. $(W_p) \Rightarrow (W_{p'})$ or $(G_{p'}) \Rightarrow (G_p)$
for $p' > p$
(OK if X: Riem., $P = P_t$)

(ii) When "$(W_p) \Rightarrow$ (pathwise control)"
in the case $P = P_t$?

(iii) "Bakry-Émery's Γ_2-criterion $\Leftrightarrow (G_q)$"
in the case $P = P_t$, $\tilde{d} = \mathrm{e}^{-kt} d$?
(When $|\nabla_d f| = |\Gamma f|^{1/2}$?).

(iv) Relation with other "lower curvature bounds"...