Duality on gradient estimates and Wasserstein controls

Kazumasa Kuwada

(Ochanomizu University

Universität Bonn)

§0 Motivation

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

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

Our goal:

Generalization of (ii) $\Leftrightarrow$ (iii), to obtain a (ii)/(iii)-type estimate from the other one.

(ii) $d_p^W(P_t^*\mu, P_t^*\nu) \le \mathrm{e}^{-kt} d_p^W(\mu,\nu)$
(L^p-Wasserstein control)

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

§1 Framework and Main Result

Framework

- (X, d): Polish, proper length metric space.

$$\Rightarrow \left(\begin{array}{l} \forall x, y \in X, \\ \exists \text{a minimal geodesic joining } x \text{ and } y \end{array} \right)$$

- $\mathcal{P}(X)$: probability measures on X.

For $\mu, \nu \in \mathcal{P}(X)$, $\Pi(\mu, \nu) \subset \mathcal{P}(X \times X)$ by

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

<u>L^p-Wasserstein distance</u>

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

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

Gradient

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

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

⋆ $|\nabla_d f|$ is an upper gradient of f,

i.e. for any curve γ joining x and y,

$$f(y) - f(x) \le \int_0^{d(x,y)} |\nabla_d f|(\gamma(s))ds.$$

- $(P_x)_{x \in X} \subset \mathcal{P}(X)$ s.t. $x \mapsto P_x$ continuous,
  $P : C_b(X) \to C_b(X)$,

$$Pf(x) := \int_M f \, dP_x .$$

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

- $\tilde{d}$: continuous geodesic distance on X.
  (e.g. $\tilde{d} = \mathrm{e}^{-kt} d$)
  We also use notations $|\nabla_{\tilde{d}} f|$ and $\tilde{d}_p^W(\mu, \nu)$.

L^p-Wasserstein control

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

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 $f \in C_b^{\mathrm{Lip}}(X)$,

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

for $q = \infty$.

Assumptions

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

(i) (X, d, v) satisfies

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

(ii) $P_x \ll v$ and $x \mapsto P_x(y)$: continuous.

Assumptions

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

(i) (X, d, v) satisfies

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

(for employing a general theory of the Hamilton-Jacobi semigroup)

(ii) $P_x \ll v$ and $x \mapsto P_x(y)$: continuous.

(technical)

Local volume doubling condition

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

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

$(1, \rho)$-local Poincaré inequality

$\forall R > 0$, $\exists \lambda \ge 1$, $\exists C_P > 0$ s.t. $\forall r < R$,

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

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

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

Theorem (K.)

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

(i) $(C_p) \Rightarrow (G_q)$.

(ii) Under Assumption (i)(ii), $(G_q) \Rightarrow (C_p)$.

Remarks

- For $p' > p$, $\begin{cases} (G_p) \Rightarrow (G_{p'}), \\ (C_{p'}) \Rightarrow (C_p). \end{cases}$
  (without Assumption (i)(ii))

- $(G_\infty) \Leftrightarrow (C_1)$ is well-known.
  (via Kantorovich-Rubinstein formula;
  without Assumption (i)(ii))

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

§2 Examples and Applications

In this §,

$P = P_t$: heat semigroup associated with $\boldsymbol{\Delta}$ (or a diffusion semigroup).

(A) How do we obtain (C_p)?
(mainly the case $\tilde{d} = \mathrm{e}^{-kt} d$)

(B) Analytic approach to (G_q).
(the case $\tilde{d} = \mathrm{e}^{-kt} d$ and beyond it)

(C) (C_p) for hypoelliptic diffusions on a Lie group.

(A) How do we obtain (C_p)?

(1) Coupling method

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

$\exists (B_t^{(1)}, B_t^{(2)})$: "infinitesimally parallel" coupling of two Brownian motions s.t.

$$d(B_t^{(1)}, B_t^{(2)}) \leq \mathrm{e}^{-kt/2} d(B_0^{(1)}, B_0^{(2)})$$

$\forall t > 0$, almost surely.

$\Rightarrow (C_\infty)$

Brownian motion under backward (super)Ricci flow

$(X, g(t))_{t\in[0,T]}$: cpl. Riem. mfds.

$\partial_t g(t) \le \mathrm{Ric}_{g(t)}$.

$(B_t)_{t\in[0,T]}$: $g(t)$-Brownian motion

(sol. to the martingale problem of $\partial_t + \frac{1}{2}\Delta_{g(\cdot)}$).

$\exists(B_t^{(1)}, B_t^{(2)})$: a coupling of two $g(t)$-BMs s.t.

$$d_{g(t)}(B_t^{(1)}, B_t^{(2)}) \le d_{g(0)}(B_0^{(1)}, B_0^{(2)})$$

(McCann & Topping '08 for (C_2).
Arnaudon & Coulibaly & Thalmaier '09,
cf. K.-Philipowski '09 for non-explosion of B_t.)

(A) How do we obtain (C_p) ?

(2) Gradient flow formulation of the heat flow

- Heuristically, heat distributions $(\mu_t)_{t\geq 0}$ is a "gradient flow" of the relative Entropy functional

$$E(\mu) = \int_X \frac{d\mu}{dv} \log \frac{d\mu}{dv}\, dv.$$

on L^2-Wasserstein space $\mathcal{P}_2(X) \subset \mathcal{P}(X)$.
(Otto '98, Ambrosio & Gigli & Savaré '05,...)

- When X: cpl. Riem. mfd,

$\mathrm{Ric} \geq k$
$\Leftrightarrow$ "$\mathrm{Hess}\, E \geq k$" on $(\mathcal{P}_2(X), d_2^W)$.

(von Renesse & Sturm '05)

Heuristically,

$$\text{“Hess } E \geq k\text{”} \quad \Rightarrow d_2^W(\mu_t, \nu_t) \leq e^{-kt} d_2^W(\mu_0, \nu_0).$$

If X: cpl. Riem. mfd., it can be made rigorous.

Moreover, $\mu_t = P_t^* \mu_0$ (identification).

(Villani '08, Erbar '08, Ohta '09,...)

$$\Rightarrow (C_2)$$

More singular spaces (e.g. Alexandrov spaces):

Savaré '07, Ohta '09,...

Remark on (A)

To obtain (C_p), we have used a notion of lower curvature bound which is different from (G_q).

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

$$
\begin{aligned}
(G_\infty) &\Rightarrow \mathbf{Ric} \geq \boldsymbol{k} \\
&\Rightarrow (C_\infty) \quad \text{(coupling method)} \\
&\Rightarrow (G_1) \quad \text{(easy)} \\
&\Rightarrow (G_\infty) \quad \text{(monotonicity)}
\end{aligned}
$$

(B) Analytic approach to (G_q)

(1) Bakry & Émery's Γ_2-criterion

A: (L^2-)generator of P_t.

- $\Gamma(f, f) := \frac{1}{2}(A(f^2) - 2fAf)$.
- $|\nabla f| := \Gamma(f, f)^{1/2}$.
- $\Gamma_2(f, f) := \frac{1}{2}(A\Gamma(f, f) - 2\Gamma(f, Af))$.

If X: cpl. Riem. mfd. & $A = \Delta$,

$$\mathrm{Ric} \geq k \Leftrightarrow \Gamma_2(f, f) \geq k\Gamma(f, f).$$

$$\Gamma_2(f,f) \geq k\Gamma(f,f) \Leftrightarrow |\nabla P_t f|(x) \leq \mathrm{e}^{-kt} P_t |\nabla f|(x)$$

$\Rightarrow$ if $|\nabla f| = |\nabla_d f|$,

$$\Gamma_2(f,f) \geq k\Gamma(f,f) \Leftrightarrow (G_1)$$

(with $\tilde{d} = \mathrm{e}^{-kt} d$).

<u>Remark</u>

$|\nabla f| = |\nabla_d f|$ does not seem to be trivial even if there is a strong connection between d and P_t.

(B) Analytic approach to (G_q)

(2) Hörmander-type operators on a Lie group

X: Lie group with a right-Haar measure v.

$\{X_i\}_{i=1}^n$: left-invariant, linearly independent vector fields generating all left-invariant vector fields in the sense of Lie algebra (Hörmander condition).

$A := \sum_{i=1}^{m} X_i^2$, P_t: semigroup associated with A.

L^q-Gradient estimate

$$|\nabla P_t f|(x) \leq K_q(t) P_t(|\nabla f|^q)(x)^{1/q}. \quad (G_q^*)$$

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 et al. '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 '09.
- $X = SU(2)$, $q > 1$, $K_q(t) = K_q e^{-t}$: Baudoin & Bonnefont '09.

"Lower curvature bound" on the Heisenberg group

- Formally, Ric / Γ_2 is unbounded from below.
- Curvature-dimension condition CD(k, N) ("$\dim X \leq N$ and $\mathrm{Ric} \geq k$") does not hold for any $N \in [1, \infty)$, $k \in \mathbb{R}$ (Juillet '09).
  - CD$(k, \infty) \Leftrightarrow$ "$\mathrm{Hess}\, E \geq k$" .
  - Weaker condition (MCP (k, N)) holds for $N = 5$ and $k = 0$ (Juillet '09).

An appropriate notion of a lower curvature bound is not clear until now!

(C) (C_p) for hypoelliptic diffusions on a Lie group

(1) General result

$X, v, \{X_i\}_{i=1}^m, P$: as before.

Carnot-Caratheodory distance

For $V \in T_x X$,

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

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

Proposition

$(\boldsymbol{X}, \boldsymbol{d}, \boldsymbol{v})$, $\boldsymbol{P} = \boldsymbol{P_t}$: as above.

(i) $(\boldsymbol{X}, \boldsymbol{d}, \boldsymbol{v}; \boldsymbol{P})$ satisfies Assumption (i)(ii).

(ii) $(\boldsymbol{G_q^*}) \Rightarrow (\boldsymbol{G_q})$.

Corollary

$(\boldsymbol{G_q^*}) \Rightarrow (\boldsymbol{C_p})$ for $\boldsymbol{q} \in [\mathbf{1}, \boldsymbol{\infty}]$.

(C) (C_p) for hypoelliptic diffusions on a Lie group

(2) 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).

(C_∞): For each $t>0$, $\exists$ a coupling $(\boldsymbol{B}_t, \tilde{\boldsymbol{B}}_t)$ of (B_t^1, B_t^2, B_t^3) with initial conditions $\boldsymbol{a} \in \mathbb{R}^3$ and $\boldsymbol{b} \in \mathbb{R}^3$ respectively s.t.

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

In this case, $\exists C_1, C_2 > 0$ s.t.

$$C_1 \|\boldsymbol{b}^{-1}\boldsymbol{a}\| \leq d(\boldsymbol{a}, \boldsymbol{b}) \leq C_2 \|\boldsymbol{b}^{-1}\boldsymbol{a}\|,$$

where $\|(x, y, z)\| = \left((x^2+y^2)^2 + z^2\right)^{1/4}$.

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 $[\mathcal{V},\mathcal{V}] = \mathcal{Z}$, $[\mathcal{V},\mathcal{Z}] = [\mathcal{Z},\mathcal{Z}] = 0$.
- $J \; : \; \mathcal{Z} \longrightarrow \mathrm{End}\,\mathcal{V}$ given by

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

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

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

Remarks

- $\forall m$, the $(2m+1)$-dim. Heisenberg group is of type H.
- A free nilpotent Lie group of step 2 is of type H iff it is the 3-dim. Heisenberg group.
- Possible dimension of a group of type H is completely determined.

§3 Sketch of the Proof

Recall:

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

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

- The case $p = 1$ ($q = \infty$) and $(C_\infty) \Rightarrow (G_1)$ are easy and well-known.

- $d_p^W(\mu,\nu) \overset{p\to\infty}{\longrightarrow} d_\infty^W(\mu,\nu) \in [0,\infty]$.
  $\Rightarrow$ For $(G_q) \Rightarrow (C_p)$, we may assume $p < \infty$.

- (C_p) for $\mu = \delta_x, \nu = \delta_y \Rightarrow (C_p)$.
  (disintegration of measures)

(A) Proof of $(C_p) \Rightarrow (\boldsymbol{G_q})$

$$\pi \in \Pi(P_x, P_y),\ \|d\|_{L^p(\pi)} = d_p^W(P_x, P_y).$$

$$\begin{aligned} Pf(x) - Pf(y) &= \int_X f\, dP_x - \int_X f\, dP_y \\ &= \int_{X \times X} (f(z) - f(w))\pi(dzdw) \\ &= \int_{\{d(z,w) \le r\}} \cdots + \int_{\{d(z,w) > r\}} \cdots \\ &= (I) + (II). \end{aligned}$$

$$(I) \leq \int_{\{d(z,w)\leq r\}} \frac{f(z)-f(w)}{d(z,w)} d(z,w)\pi(dzdw)$$

$$\leq \left\{ \int_X \sup_{w\in B_r(z)} \left| \frac{f(z)-f(w)}{d(z,w)} \right|^q P_x(dz) \right\}^{1/q}$$

$$\times\, d_p^W(P_x, P_y)$$

$$\underset{(C_p)}{\boxed{\leq}} \underbrace{\{\cdots\}^{1/q}}_{\substack{\big\downarrow r\to 0 \\ P(|\nabla_d f|^q)(x)^{1/q}}} \tilde{d}(x,y).$$

$$(II) \leq 2\|f\|_\infty \int_{\{d(z,w)>r\}} \pi(dzdw)$$

$$\text{Chebyshev} \boxed{\leq} 2\|f\|_\infty \frac{d_p^W(P_x, P_y)^p}{r^p}$$

$$(C_p) \boxed{\leq} 2\|f\|_\infty \tilde{d}(x,y) \frac{\tilde{d}(x,y)^{p-1}}{r^p}.$$

Choose $r = r(x,y)$ s.t.

$$\boxed{\lim_{y\to x} r = 0 \ \& \ \lim_{y\to x} \frac{\tilde{d}(x,y)^{p-1}}{r^p} = 0}$$

(e.g. $r = \tilde{d}(x,y)^{1/(2q)}$). ■

(B) Proof of $(G_q) \Rightarrow (C_p)$

General theory of the Hamilton-Jacobi semigroup

(Lott & Villani '07, Balogh et al. '09)

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

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

$$\partial_t Q_t f(x) = -\frac{1}{q} |\nabla_d Q_t f|(x)^q \ .$$

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

Kantorovich duality

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

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

$$= p\, Q_1(p^{-1} f)(x).$$

$$\Downarrow$$

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

$$\begin{cases} \gamma : [0,1] \to X \quad \tilde{d}\text{-minimal geodesic}, \\ \gamma_0 = y, \ \gamma_1 = x, \\ \tilde{d}(\gamma_s, \gamma_t) = |t - s|\tilde{d}(x, y). \end{cases}$$

—— ∘ —— ∘ ——

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

$$\text{“ = ”} \quad \sup_f \left[\int_0^1 \partial_t (PQ_t f(\gamma_t)) dt \right].$$

interpolation

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

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

Questions

(i) When does $(C_p) \Rightarrow (C_{p'})$ / $(G_{p'}) \Rightarrow (G_p)$ occur for $p' > p$?
(OK if X: Riem., $P = P_t$)

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

(iii) When does $|\nabla_d f| = |\nabla f| (= \Gamma(f, f)^{1/2})$ hold?
(OK if X: (sub-)Riem., $P = P_t$)

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