401-3382-25L: Log-Sobolev Inequalities and Markov Semigroups

Lecture 4 – Bakry-Émery Criterion

Lecturer: Yuansi Chen *Week 8-9* *Spring 2025*

Key concepts:

- Carré du champ, iterated carré du champ
- Curvature - dimension condition
- Gamma-calculus to derive Poincaré inequality
- Langevin semigroup
- Weak and strong commutation
- Log-Sobolev inequality

4.1 Gamma calculus

Let (X_t) be a Markov process, E be the state space, (P_t) be the associated semigroup and L be the generator. Given a test function f, and $t > 0$, we consider

$$\beta(s) := P_s\left((P_{t-s}f)^2\right), s \in [0,t]. \tag{4.1}$$

This is the same quantity that we have seen in the proof of local Poincaré inequality (in Week 3-4).

In the Gaussian Poincaré inequality case, we have seen that taking derivative of $\beta(s)$ reduces the Poincaré inequality to bounding terms along the derivative, which we need to use the so-called sub-commutation. The main purpose of the first part of this lecture is to check whether the same idea to more general measures.

Notice that

$$\beta(t) - \beta(0) = P_t(f^2) - (P_t f)^2.$$

So if the semigroup P_t admits a stationarity measure μ (i.e., $\int P_t f d\mu = \int f d\mu, \forall f$) and μ is ergodic (i.e., $P_t f \to \int f d\mu$), then letting $t \to \infty$, the LHS of the above equation becomes $\mathrm{Var}_\mu(f)$, which is the term in Poincaré inequality we want to upper bound. To upper bound this quantity, we analyze its derivative. We start by establishing a few properties of β.

Lemma 1. *We have*

(a) *β is non-decreasing. That is,*

$$\beta(s+\epsilon) \geq \beta(s), \quad \forall s \geq 0, \epsilon > 0$$

(b) *The first derivative*

$$\beta'(s) = 2P_s(\Gamma(g)),$$

where $g = P_{t-s}f$, and $\Gamma(g) := \frac{1}{2}\left[Lg^2 - 2gLg\right]$ is the ***carré du champ****. More generally, the carré du champ is introduced via the bilinear form*

$$\Gamma(h_1, h_2) := \frac{1}{2}\left[L(h_1 h_2) - h_1(Lh_2) - (Lh_1)h_2\right],$$

and $\Gamma(h) = \Gamma(h,h) \geq 0$ because β is nondecreasing.

(c) *The second derivative*

$$\begin{aligned}\beta''(s) &= 2P_s(L\Gamma(g) - 2\Gamma(Lg, g)) \\ &= 4P_s\left(\Gamma_2(g)\right),\end{aligned}$$

where $\Gamma_2(h) := \frac{1}{2}\left[L\Gamma(h) - 2\Gamma(Lh, h)\right]$ is called the ***iterated version of the carré du champ****.*

Proof of Lemma 1. (a) Using the semigroup property, we have

$$\begin{aligned}\beta(s+\epsilon) &= P_s P_\epsilon\left((P_{t-s-\epsilon}f)^2\right) \\ &\overset{(i)}{\geq} P_s\left((P_\epsilon P_{t-s-\epsilon}f)^2\right) \\ &= P_s\left((P_{t-s}f)^2\right) = \beta(s),\end{aligned}$$

(i) follows from Jensen's inequality applied to the convex function $x \mapsto x^2$.

(b) Recall that using the definition of the infinitesimal generator L, $\partial_s P_s = LP_s = P_s L$. Then with $g = P_{t-s}f$,

$$\begin{aligned}\beta'(s) &= \partial_s P_s(g^2) + P_s(2(\partial_s g)g) \\ &= LP_s(g^2) + P_s(2(-Lg)g) \\ &= P_s\left[Lg^2 - 2gLg\right] \\ &= 2P_s\Gamma(g).\end{aligned}$$

To prove that $\Gamma(h) \geq 0$, we notice that

$$\beta'(s) = 2P_s(\Gamma(P_{t-s}f)).$$

Let $s = 0$, using that β is nondecreasing, $\beta'(0) = \Gamma(P_t f) \geq 0$. Finally, let $t \to 0$, yields the nonnegativity of Γ.

(c) First, using that Γ is a bilinear form, we have

$$\partial_s \Gamma(g) = 2\Gamma(\partial_s g, g) = -2\Gamma(Lg, g).$$

We obtain

$$\begin{aligned}\beta''(s) &= 2(\partial_s P_s)(g) + 2P_s(\partial_s \Gamma(g)) \\ &= 2P_s\left[L\Gamma(g) - 2\Gamma(Lg, g)\right] \\ &= 4P_s\Gamma_2(g).\end{aligned}$$

□

Intuitively, if we have a way to relate $\beta''(s)$ with $\beta'(s)$, then we can bound $\beta'(s)$, which in turn would allow us to bound the desired quantity $\beta(t) - \beta(0)$. Let's recall what happened when (X_t) is the Ornstein-Uhlenbeck (OU) semigroup.

Example 1 (β in OU semigroup)**.** *When (X_t) is the OU semigroup, we know the infinitesimal generator L takes the form*

$$Lf = \Delta f - \langle x, \nabla f \rangle .$$

Having the expression for L is sufficient to derive both Γ and Γ_2. After some basic derivative calculation, assuming f is twice differentiable, we obtain

$$\begin{aligned}\Gamma(f) &= |\nabla f|^2 \\ \Gamma_2(f) &= \sum_{ij} (\partial_{ij} f)^2 + |\nabla f|^2 .\end{aligned}$$

Then we always have $\Gamma_2(g) \geq \Gamma(f)$, which implies that

$$\beta''(s) \geq 2\beta'(s).$$

Solving the differential inequality via Grönwall's lemma, we obtain

$$\begin{aligned}&\beta'(s) \geq e^{2s}\beta'(0), \text{ and} \\ &\beta'(s) \leq e^{-2(t-s)}\beta'(t) = 2e^{-2(t-s)}P_t(\Gamma(f)).\end{aligned}$$

With the second inequality, we have

$$\beta(t) - \beta(0) = \int_0^t \beta'(s)ds \leq 2P_t\Gamma(f)\int_0^t e^{-2(t-s)}ds = 2P_t\Gamma(f)\frac{1-e^{-2t}}{2} = \left(1-e^{-2t}\right)P_t(|\nabla f|^2).$$

Letting $t \to \infty$, we obtain the Gaussian Poincaré inequality.

Remark 1. *The inequality $\beta'(t) \geq e^{2t}\beta'(0)$ is exactly*

$$|\nabla P_t f|^2 \leq e^{-2t} P_t(|\nabla f|^2).$$

In Week 3-4, we obtained this sub-commutation property using the explicit formula of the OU semigroup, via the Mehler formula. The proof only uses the relationship between β' and β''. This is exactly the purpose of Gamma-calculus, we want to do similar things for semigroups without an explicit expression.

We introduce one sufficient condition to relate β' and β''.

Curvature-dimension condition. We say that the Markov process (X_t) satisfies the *curvature-dimension condition* $\mathrm{CD}(\rho, n)$, if for any function f we have

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{1}{n}(Lf)^2.$$

Remark 2. *Since we are mostly interested in dimension-free functional inequalities, in this course, we focus on the curvature-dimension condition $CD(\rho, \infty)$, which reads*

$$\Gamma_2(f) \geq \rho\Gamma(f).$$

The following establishes the equivalence between weak commutation and the curvature-dimension condition.

Lemma 2 (Weak commutation)**.** *Let $\rho > 0$, the following two are equivalent*

(a) The semigroup (P_t) satisfies $CD(\rho, \infty)$

(b) For every function f and $t > 0$, we have $\Gamma(P_t f) \leq e^{-2\rho t} P_t \Gamma f$.

Proof of Lemma 2. We have

$$\beta'(s) = 2P_s(\Gamma(P_{t-s}f))$$
$$\beta''(s) = 4P_s(\Gamma_2(P_{t-s}f)).$$

"$(a) \implies (b)$." $\mathrm{CD}(\rho, \infty)$ implies that $\beta''(s) \geq 2\rho\beta'(s)$. Grönwall's lemma gives (2).
"$(b) \implies (a)$." (2) implies that $\beta'(s) \geq e^{2\rho s}\beta'(0)$. Taylor expansion around $s = 0$, gives

$$\beta''(0) \geq 2\rho\beta'(0),$$

which is $\Gamma_2(P_t f) \geq \rho\Gamma(P_t f)$. Let $t \to 0$ to obtain $\mathrm{CD}(\rho, \infty)$. □

4.2 Poincaré inequality from curvature-dimension condition

We first define the general version of Poincaré inequality. We assume that the semigroup P_t admits a stationarity measure μ (i.e., $\int P_t f d\mu = \int f d\mu, \forall f$) and μ is ergodic (i.e., $P_t f \to \int f d\mu$).

Dirichlet form The *Dirichlet form* is the quadratic form given by

$$\mathcal{E}(f,g) := \int_E \Gamma(f,g) d\mu.$$

And we write $\mathcal{E}(f) := \mathcal{E}(f,f)$. Note that by stationarity, we have $\int L(f^2) d\mu = 0$ and the Dirichlet has a simplified form

$$\mathcal{E}(f) = \int \Gamma(f) d\mu = -\int f(Lf) d\mu.$$

General Poincaré inequality We say μ satisfies the Poincaré inequality with constant C_P if for every test function f,

$$\mathrm{Var}_\mu(f) \leq C_P \mathcal{E}(f).$$

Theorem 4.2.1 (Poincaré via curvature-dimension condition (or Gamma-two criterion, or Bakry-Émery criterion))**.** *For $\rho > 0$, $CD(\rho, \infty)$ implies that μ satisfies Poincaré inequality with constant $1/\rho$.*

Proof of Theorem 4.2.1. For $\beta(s) = P_s((P_{t-s}f)^2)$ we have

$$\beta'(s) = 2P_s(\Gamma P_{t-s} f).$$

Using the weak commutation in Lemma 2, we obtain

$$\beta'(s) \leq e^{-2\rho(t-s)} \beta'(t) = e^{-2\rho(t-s)} \cdot 2P_t(\Gamma f).$$

Therefore,

$$P_t(f^2) - (P_t f)^2 = \beta(t) - \beta(0) = \int_0^t \beta'(s) ds \leq \frac{1 - e^{-2\rho t}}{2\rho} 2P_t(\Gamma(f)).$$

Letting $t \to \infty$, and using ergodicity, we conclude. $\square$

Just like in the Gaussian Poincaré inequality case, the Poincaré inequality is equivalent to an exponential decay of variance.

Proposition 4.2.1 (Variance decay)**.** *The following two statements are equivalent:*

(a) μ satisfies the Poincaré inequality with constant C_P.

(b) For any function f, $\mathrm{Var}_\mu(P_t f) \le e^{-2t/C_P}\,\mathrm{Var}_\mu(f)$.

Proof. Note that

$$\frac{d}{dt}\mathrm{Var}(P_t f) = \frac{d}{dt}\int (P_t f)^2 d\mu = -2\mathcal{E}(P_t f).$$

"$(a) \implies (b)$", applying Poincaré allows to bound RHS by variance again. Then it suffices use Grönwall's lemma.
"$(b) \implies (a)$", Taylor expansion of $\mathrm{Var}_\mu(P_t f) \le e^{-2t/C_P}\,\mathrm{Var}_\mu(f)$ around $t = 0$, and take limit $t \to 0$. □

Next, we show that if μ is reversible, then the Poincaré inequality is equivalent to an integrated version of the $\mathrm{CD}(\rho, \infty)$ condition.

Proposition 4.2.2 (Integrated Gamma-two criterion)**.** *Let $\rho > 0$. Consider the following two:*

(a) For any f, we have

$$\int_E \Gamma_2(f)d\mu \ge \rho \int_E \Gamma(f)d\mu.$$

(b) μ satisfies Poincaré with constant $1/\rho$.

We always have $(a) \implies (b)$ and if μ is reversible, they are equivalent.

Proof. **TODO** □

4.3 Langevin semigroup

Given a measure μ on $\mathbb{R}^n$, to prove a Poincaré inequality using the above Gamma-calculus, we first need to find a semigroup which has its stationary measure μ. Can we always find a semigroup?

Given $\mu = \frac{1}{Z}e^{-V}$, where $Z = \int_{\mathbb{R}^n} e^{-V}$ is the normalization constant. Assume $V : \mathbb{R}^n \to \mathbb{R}$ is a smooth function. Consider the Langevin stochastic differential equation (SDE)

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dB_t.$$

If ∇V is Lipschitz, then the above SDE has a unique strong solution (see e.g. [Oks13]). Then its solution (X_t) is a Markov process and in this section, let (P_t) be the associated semigroup.

Langevin semigroup We have

1. $P_t f(x) = \mathbb{E}[f(X_t) \mid X_0 = x]$

2. Its infinitesimal generator is

$$Lf = \Delta f - \langle \nabla V, \nabla f \rangle .$$

3. μ is reversible for the process.

4. μ is ergodic.

Proof. **TODO** □

Remark 3. *When $V = |x|^2/2$, the semigroup (P_t) is the Ornstein-Uhlenbeck semigroup and the stationary measure μ is the standard Gaussian. The Langevin semigroup can be seen as a generalization. The main difficulty of the generalization is that (P_t) no longer has an explicit expression any more. We hope to impose sufficient conditions on V such that the Gamma-calculus applies, and then Poincaré inequality and log-Sobolev inequality can be obtained as we did for the OU semigroup.*

Lemma 3 (Gamma-calculus for Langevin semigroup)**.** *Given the infinitesimal generator $Lf = \Delta f - \langle \nabla V, \nabla f \rangle$, we have*

$$\begin{aligned} \Gamma(f, g) &= \langle \nabla f, \nabla g \rangle \\ \Gamma_2(f) &= \operatorname{trace}\left((\nabla^2 f)^2\right) + \left\langle \nabla^2 V \nabla f, \nabla f \right\rangle . \end{aligned}$$

Here $\operatorname{trace}((\nabla^2 f)^2) = \sum_{ij} (\partial_{ij} f)^2$.

Lemma 4 (When Langevin semigroup satisfies CD)**.** *Let $\rho > 0$. The Langevin semigroup satisfies $CD(\rho, \infty)$ if and only if*

$$\nabla^2 V \succeq \rho \mathbb{I}_n .$$

Proof. If $\nabla^2 V \succeq \rho \mathbb{I}_n$, using the expression of Γ_2 in Lemma 3, we have

$$\Gamma_2(f) \geq \rho |\nabla f|^2 = \rho \Gamma(f).$$

Conversely, taking linear functions $f(x) = \langle u, x \rangle$, then $\nabla^2 f = 0$ and $\nabla f = u$. From CD, we have

$$\left\langle \nabla^2 V u, u \right\rangle \geq \rho \langle u, u \rangle .$$

It holds for all u, we conclude. □

Corollary 1. *Let $\rho > 0$, and let $\mu = e^{-V}$ is a probability measure. Assume $\nabla^2 V \succeq \rho \mathbb{I}_n$ pointwise. Then μ satisfies the Poincaré inequality*

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\rho} \int_{\mathbb{R}^n} |\nabla f|^2 \, d\mu.$$

Proof. Apply Lemma 4 and then Theorem 4.2.1. □

Remark 4. *It is possible to generalize the above argument (Bakry-Émery criterion) to a Riemmanian manifold $\mathcal{M}$ equipped with the normalized volume measure vol_g, the Laplace-Beltrami operator, and a vector field. Bochner's formula gives the form of Γ_2 for Riemmanian Langevin semigroup, where a new term $Ric(\nabla, \nabla)$ appears in Γ_2. And the Bakry-Émery curvature in the curvature-dimension condition becomes the sum of a curvature of the space and a curvature of the Markov process.*

4.4 Gamma-calculus for general ϕ-entropy of a diffusion

We have seen how the Gamma calculus enables us to derive Poincaré inequality. We would like to do the same for log-Sobolev inequality. Recall the notion of ϕ-entropy when we derived Gaussian log-Sobolev inequality. Given a positive function f, we are interested in

$$\beta(s) = P_s\left(\phi\left(P_{t-s} f\right)\right),$$

where $\phi = x \log x$. Notice that

$$\beta(t) - \beta(0) = P_t(\phi(f)) - \phi(P_t(f)).$$

So if the semigroup P_t admits a stationary measure μ and μ is ergodic, then letting $t \to \infty$, the above expression recovers $\mathrm{Ent}_\mu(f)$, which is the term in log-Sobolev inequality we want to upper bound.

(Continuous) diffusion We say a Markov process (X_t) is a *diffusion* if its infinitesimal generator L satisfies

$$L(\phi(f)) = \phi'(f) L f + \phi''(f) \Gamma(f).$$

Typically, a diffusion is the solution of a stochastic differential equation driven by the Brownian motion.

Example 2 (Langevin semigroup is a diffusion)**.** *For the Langevin semigroup, we have*

$$Lf = \Delta f - \langle \nabla V, \nabla f \rangle .$$

And

$$\begin{aligned} L\phi f &= \Delta(\phi(f)) - \langle \nabla V, \nabla \phi(g) \rangle \\ &= \phi'(f) Lf + \phi''(f) \left|\nabla f\right|^2 . \end{aligned}$$

Lemma 5. *We have*

(a) β is non-decreasing. That is,

$$\beta(s+\epsilon) \geq \beta(s), \quad \forall s \geq 0, \epsilon > 0$$

(b) The first derivative

$$\beta'(s) = P_s \left[L(\phi(g)) - \phi'(g) Lg \right]$$

where $g = P_{t-s} f$. In the case of diffusion, it simplifies to

$$\begin{aligned} \beta'(s) &= P_s(\phi''(g)\Gamma(g)) \\ &\overset{(i)}{=} P_s(\Gamma(g)/g) \end{aligned}$$

(i) follows when $\phi = x \log x, \phi'' = 1/x$.

Proof. (a) It follows the same proof using Jensen's inequality and ϕ is convex.

(b) When taking derivative, it hits s in two places.

□

Remark 5. *The main reason that we focus on (continuous) diffusion is that we can simplify the first derivative β'. According to the above lemma, normally, the term Lg in β' makes it look like that β' depends on the second derivative of g. However, as it was the case in Lemma 1, β' only depends on Γ, which only depends on the first derivative of g, when (P_t) is a continuous diffusion. Intuitively, having a continuous diffusion allows us to have integration by parts, and integration by parts saves us a derivative!*

A typical example of a Markov that is not a continuous diffusion is a process taking values on a discrete space, like the random walk on a graph or on the Boolean hypercube $\{-1,1\}^n$.

Lemma 6 (Strong commutation)**.** *If (P_t) is a diffusion satisfying $CD(\rho, \infty)$ then*

$$\sqrt{\Gamma P_t f} \leq e^{-\rho t} P_t \left(\sqrt{\Gamma f} \right) .$$

Proof. Let $\alpha(s) = P_s\left(\sqrt{\Gamma(P_{t-s}f)}\right)$. We have

$$\alpha'(s) = P_s\left(\Gamma(g)^{-1/2}\Gamma_2(g) - \frac{1}{4}g^{-3/2}\Gamma\Gamma g\right).$$

Then

$$\alpha'(s) - \rho\alpha(s) = P_s\left(\Gamma(g)^{-1/2}\left(\Gamma_2(g) - \rho\Gamma(g) - \frac{\Gamma(\Gamma(g))}{4\Gamma(g)}\right)\right).$$

We want to show that $\alpha'(s) \geq \rho\alpha(s)$ and then complete the proof via Grönwall. It suffices to show that

$$\Gamma_2(g) \geq \rho\Gamma(g) + \frac{\Gamma(\Gamma(g))}{4\Gamma(g)},$$

which will be addressed in the next Lemma. □

Lemma 7. *If (P_t) is a diffusion satisfying $CD(\rho,\infty)$ then*

$$\Gamma_2(f) \geq \rho\Gamma(f) + \frac{\Gamma(\Gamma(f))}{4\Gamma(f)},$$

for all f.

Proof. **TODO** □

4.5 Log-Sobolev inequality for a diffusion

We are ready to prove the log-Sobolev inequality for a diffusion.

Theorem 4.5.1. *If (P_t) is a diffusion satisfying $CD(\rho,\infty)$ for some $\rho > 0$ and has an ergodic stationary measure μ. Then μ satisfies the following logarithmic Sobolev inequality, for any positive f,*

$$Ent_\mu(f) \leq \frac{1}{2\rho}\int_E \frac{\Gamma(f)}{f}d\mu.$$

Remark 6. *By the diffusion property $\Gamma(f) = 4f\Gamma(\sqrt{f}) = f\Gamma(f,\log f)$. So the right-hand side in the log-Sobolev inequality can be written as*

$$\int_E \frac{\Gamma(f)}{f}d\mu = \mathcal{E}(f,\log f) = 4\mathcal{E}(\sqrt{f}).$$

Being able to interchange these three expressions is really a blessing of having a diffusion! One might find it difficult to do the same for (P_t) on a discrete space.

Corollary 2. *Let μ be a probability measure on $\mathbb{R}^n$ and $\mu = e^{-V}$. If $\nabla^2 V \succeq \rho \mathbb{I}_n$ for some $\rho > 0$, then μ satifies the following log-Sobolev inequality, for any positive f, we have*

$$Ent_\mu(f) \leq \frac{1}{2\rho} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\mu.$$

The corollary follows directly from Theorem 4.5.1 using Langevin semigroup which has $\Gamma(f) = |\nabla f|^2$ and Lemma 4.

Proof of Theorem 4.5.1. We have, for $g = P_{t-s} f$

$$\begin{aligned} \beta(s) &= P_s(g \log g) \\ \beta'(s) &= P_s \left(\Gamma(g)/g\right). \end{aligned}$$

We upper bound $\beta'(s)$ as follows

$$\begin{aligned} \frac{\Gamma(g)}{g} &= \frac{\Gamma(P_{t-s} f)}{P_{t-s} f} \\ &\overset{(i)}{\leq} e^{-2\rho(t-s)} \frac{P_{t-s}(\sqrt{\Gamma f})^2}{P_{t-s} f} \\ &\overset{(ii)}{\leq} e^{-2\rho(t-s)} P_{t-s} \left(\frac{\Gamma f}{f}\right). \end{aligned}$$

(i) follows from $\mathrm{CD}(\rho, \infty)$ and strong commutation in Lemma 6. (ii) follows from Cauchy-Schwarz inequality. Therefore,

$$\beta'(s) \leq e^{-2\rho(t-s)} P_t \left(\frac{\Gamma f}{f}\right).$$

Integrating the above from 0 to t, we obtain

$$P_t(f \log f) - (P_t f) \log(P_t f) \leq \frac{1 - e^{-2\rho t}}{2\rho} P_t \left(\frac{\Gamma f}{f}\right).$$

Letting $t \to \infty$ and using ergodicity, we conclude. □

Just as we have seen for the Gaussian LSI, for a diffusion

- LSI is equivalent to an exponential decay of entropy. **TODO**
- LSI is equivalent to hypercontractivity. **TODO**

Bibliography

[Oks13] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.