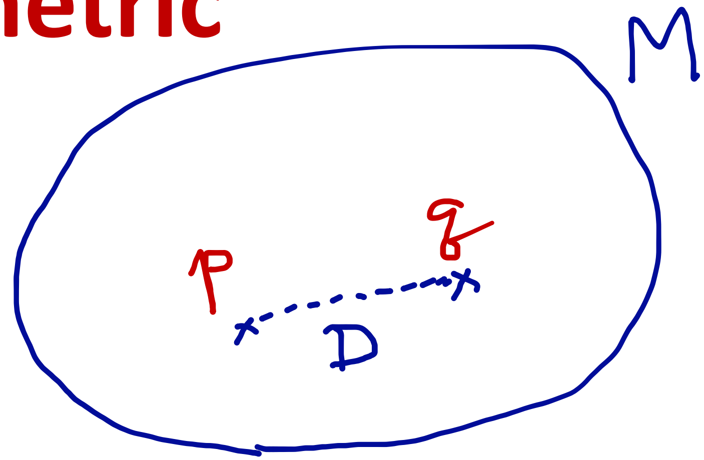
GSI – 2015 - Paris

Standard Divergence in Manifold of Dual Affine Connections

Shun-ichi Amari (RIKEN Brain Science Institute)
Nihat Ay (Max-Planck Inst. Mathematics in Science)

Divergence and metric



$$D[p:q] \geq 0$$

$$D[\boldsymbol{\xi} : \boldsymbol{\xi} + d\boldsymbol{\xi}] = \frac{1}{2} g_{ij}(\boldsymbol{\xi}) d\xi^i d\xi^j + O\left(|d\boldsymbol{\xi}|^3\right)$$

G : **Riemannian metric, positive-definite**

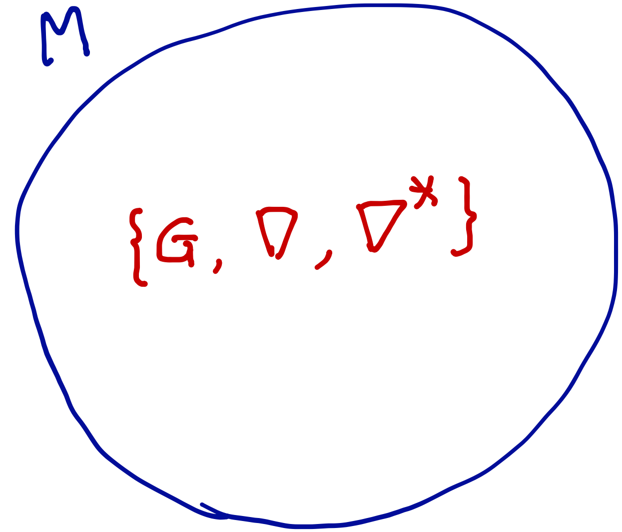
Divergence and dual affine connections

$$\Gamma_{ijk} \sim \nabla \qquad \Gamma^*_{ijk} \sim \nabla^*$$

$$\Gamma_{ijk} = -\partial_i \partial_j \partial'_k D[\xi : \xi']_{\xi'=\xi}$$

$$\Gamma^*_{ijk} = -\partial'_i \partial'_j \partial_k D[\xi : \xi']_{\xi'=\xi}$$

$$\partial_i = \frac{\partial}{\partial \xi^i}; \qquad \partial'_j = \frac{\partial}{\partial \xi'^j}$$



Dual geometry

± α duality

$$\{M, g, \nabla, \nabla^*\} \leftarrow D[p:q]$$

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla^*_X Z\rangle$$

$$\{M, g, T\}, \qquad T_{ijk} = \Gamma^*_{ijk} - \Gamma_{ijk}$$

D

$$\Gamma^{\pm\alpha}_{ijk} = \Gamma^{o}_{ijk} \mp \frac{\alpha}{2} T_{ijk}$$

Γ^{o}: **Levi-Civita connection** divergence ⟺ geometry

Dual geometry ➡ canonical divergence

M : dually flat $: \exists \psi(\boldsymbol{\theta}), \varphi(\boldsymbol{\eta})$

$$D[\boldsymbol{\theta} : \boldsymbol{\theta}'] = \psi(\boldsymbol{\theta}) + \varphi(\boldsymbol{\eta}') - \boldsymbol{\theta} \cdot \boldsymbol{\eta}'$$

Bregman divergence

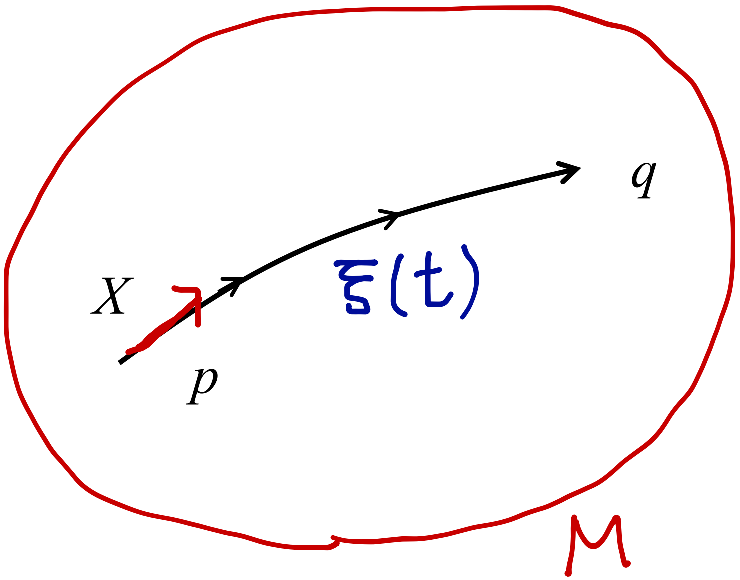
Exponential map : $\xi(t)$ geodesic

$$\nabla_{\dot{\xi}} \dot{\xi} = 0$$

$$\xi(0) = p$$

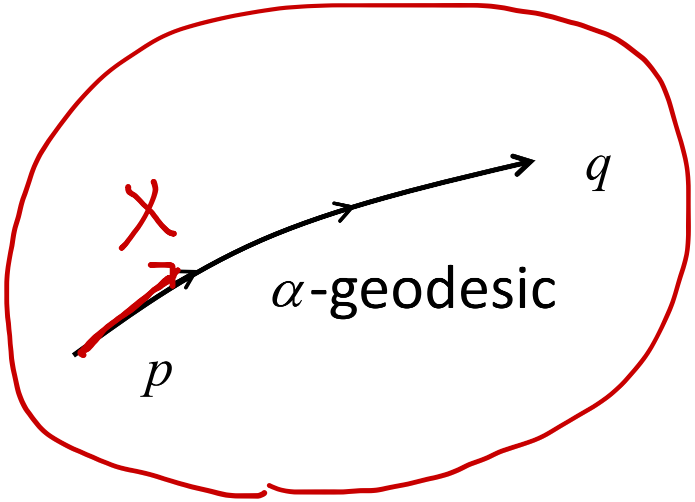
$$\xi(1) = q$$

$$\dot{\xi}(0) = X = \log_p q$$



Exponential map divergence

$$D[p:q] = \left\| X(p:q) \right\|^2$$



α-**divergence**

$$D_\alpha[p:q] = \left\| X_\alpha(p:q) \right\|^2$$

Theorem 1. Exponential map divergence induces $\alpha = -3$ geometry

Theorem 2. $\alpha = -\frac{1}{3}$ exponential map divergence recovers the original geometry

Standard divergence: $D_{\mathrm{stan}}\left[p : q\right] = \left\| X_{-1/3}\left(p, q\right) \right\|^2$

$$D_{\text{stan}}[p:q] \neq D^*_{\text{stan}}[q:p]$$

$$D_{\text{stan}}[p:q] = \frac{1}{2}\left(\left\|X_{-1/3}(p,q)\right\|^2 + \left\|X_{1/3}(q,p)\right\|^2\right)$$

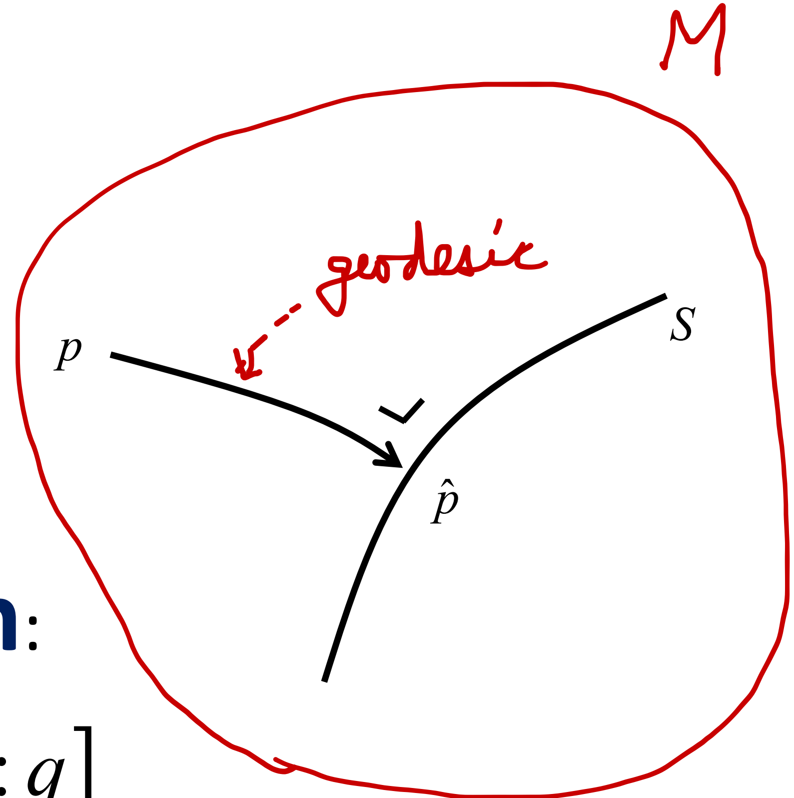
Remark: dually flat case

$$D_{\text{stan}} \neq D_{\text{can}}$$

$$D[p:q] = \int\int_0^1 t \,\|\, \dot{\xi}(t) \,\|^2 \, dt$$

Divergence and projection

$$\hat{p} = \arg\min_{q \in S} D[p:q]$$



projection theorem:

$$X = c \ \mathrm{grad}_q \ D[p:q]$$

dually flat

α-simplex