The Exponential Formula for the Wasserstein Metric

Katy Craig
UCLA

SIAM Annual Meeting, Chicago

July 8, 2014

Plan

- Gradient flow
- Discrete gradient flow
- Euler-Lagrange equation
- Exponential formula

Plan

- Gradient flow
- Discrete gradient flow
- Euler-Lagrange equation
- Exponential formula

Gradient Flow

$$\frac{\partial u(t)}{\partial t} = -\nabla E(u(t)), \quad u(0) = u$$

Gradient Flow

$$\frac{\partial u(t)}{\partial t} = -\nabla E(u(t)), \quad u(0) = u$$

Heat Equation as Gradient Flow on $L^2(\mathbb{R}^d)$

Gradient Flow

$$\frac{\partial u(t)}{\partial t} = -\nabla E(u(t)), \quad u(0) = u$$

Heat Equation as Gradient Flow on $L^2(\mathbb{R}^d)$

$L^2(\mathbb{R}^d)$ gradient:

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{E(u + hv) - E(u)}{h} \text{ for all } v \in L^2(\mathbb{R}^d)$$

Gradient Flow

$$\frac{\partial u(t)}{\partial t} = -\nabla E(u(t)), \quad u(0) = u$$

Heat Equation as Gradient Flow on $L^2(\mathbb{R}^d)$

$L^2(\mathbb{R}^d)$ gradient:

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{E(u+hv) - E(u)}{h} \text{ for all } v \in L^2(\mathbb{R}^d)$$

Thus, for $E(u) = \frac{1}{2} \int |\nabla u|^2$,

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{1}{2} \frac{\int |\nabla(u+hv)|^2 - \int |\nabla u|^2}{h} = (\nabla u, \nabla v)_{L^2} = (-\Delta u, v)_{L^2} .$$

Gradient Flow

$$\frac{\partial u(t)}{\partial t} = -\nabla E(u(t)), \quad u(0) = u$$

Heat Equation as Gradient Flow on $L^2(\mathbb{R}^d)$

$L^2(\mathbb{R}^d)$ gradient:

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{E(u + hv) - E(u)}{h} \text{ for all } v \in L^2(\mathbb{R}^d)$$

Thus, for $E(u) = \frac{1}{2} \int |\nabla u|^2$,

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{1}{2} \frac{\int |\nabla(u + hv)|^2 - \int |\nabla u|^2}{h} = (\nabla u, \nabla v)_{L^2} = (-\Delta u, v)_{L^2} \, .$$

Hence, the L^2 gradient flow of E is

$$\partial u / \partial t = -\nabla_{L^2} E(u) = -(-\Delta u) = \Delta u \, .$$

Gradient Flow

$$\frac{\partial u(t)}{\partial t} = -\nabla E(u(t)), \quad u(0) = u$$

Heat Equation as Gradient Flow on $L^2(\mathbb{R}^d)$

$L^2(\mathbb{R}^d)$ gradient:

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{E(u+hv) - E(u)}{h} \text{ for all } v \in L^2(\mathbb{R}^d)$$

Thus, for $E(u) = \frac{1}{2} \int |\nabla u|^2$,

$$(\nabla_{L^2} E(u), v)_{L^2} = \lim_{h \to 0} \frac{1}{2} \frac{\int |\nabla(u+hv)|^2 - \int |\nabla u|^2}{h} = (\nabla u, \nabla v)_{L^2} = (-\Delta u, v)_{L^2} .$$

Hence, the L^2 gradient flow of E is

$$\partial u / \partial t = -\nabla_{L^2} E(u) = -(-\Delta u) = \Delta u .$$

Note: $\nabla_{L^2} E(u) = \frac{\delta E}{\delta u}$

Examples of Hilbert Space Gradient Flow

	PDE	Energy Functional	Metric
Allen-Cahn	$\frac{d}{dt}u = \Delta u - F'(u)$	$E(u) = \frac{1}{2}\int \left[\lvert\nabla u\rvert^2 + F(u)\right]$	L^2
Cahn-Hilliard	$\frac{d}{dt}u = \Delta(\Delta u - F'(u))$	$E(u) = \frac{1}{2}\int \left[\lvert\nabla u\rvert^2 + F(u)\right]$	H^{-1}
Porous Media / Fast Diffusion	$\frac{d}{dt}u = \Delta u^m$	$E(u) = \frac{1}{m+1}\int u^{m+1}$	H^{-1}

Examples of Hilbert Space Gradient Flow

	PDE	Energy Functional	Metric
Allen-Cahn	$\frac{d}{dt}u = \Delta u - F'(u)$	$E(u) = \frac{1}{2}\int \left[\lvert\nabla u\rvert^2 + F(u)\right]$	L^2
Cahn-Hilliard	$\frac{d}{dt}u = \Delta(\Delta u - F'(u))$	$E(u) = \frac{1}{2}\int \left[\lvert\nabla u\rvert^2 + F(u)\right]$	H^{-1}
Porous Media / Fast Diffusion	$\frac{d}{dt}u = \Delta u^m$	$E(u) = \frac{1}{m+1}\int u^{m+1}$	H^{-1}

Why gradient flow?

- Free estimates, e.g. $|u(t) - v(t)| \leq e^{-\lambda t}|u(0) - v(0)|$
- Method to construct and approximate solutions (discrete gradient flow)

Wasserstein Gradient Flow

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)), \quad \mu(0) = \mu$$

Wasserstein Gradient Flow

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)), \quad \mu(0) = \mu$$

Simplifying assumptions: $\int |x|^2 d\mu < +\infty,\ \mu << \mathcal{L}^d,\ E(\mu) < +\infty$

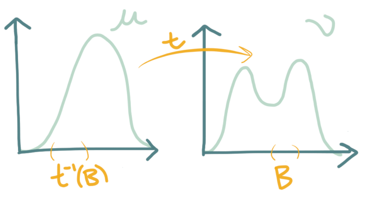
Wasserstein Gradient Flow

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)), \quad \mu(0) = \mu$$

Simplifying assumptions: $\int |x|^2 d\mu < +\infty,\ \mu << \mathcal{L}^d,\ E(\mu) < +\infty$

Wasserstein Metric

- $\mathbf{t} : \mathbb{R}^d \to \mathbb{R}^d$ transports μ to ν in case $\nu(B) = \mu(\mathbf{t}^{-1}(B))$. Write $\mathbf{t}\#\mu = \nu$.



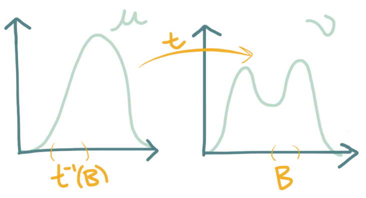
Wasserstein Gradient Flow

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)), \quad \mu(0) = \mu$$

Simplifying assumptions: $\int |x|^2 d\mu < +\infty, \ \mu << \mathcal{L}^d, \ E(\mu) < +\infty$

Wasserstein Metric

- $\mathbf{t} : \mathbb{R}^d \to \mathbb{R}^d$ transports μ to ν in case $\nu(B) = \mu(\mathbf{t}^{-1}(B))$. Write $\mathbf{t}\#\mu = \nu$.



- $W_2(\mu, \nu) = \inf \left\{ \left(\int |\mathbf{t}(x) - x|^2 d\mu \right)^{1/2} : \mathbf{t}\#\mu = \nu \right\}$

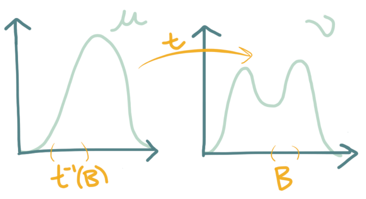
Wasserstein Gradient Flow

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)), \quad \mu(0) = \mu$$

Simplifying assumptions: $\int |x|^2 d\mu < +\infty,\ \mu << \mathcal{L}^d,\ E(\mu) < +\infty$

Wasserstein Metric

- $\mathbf{t} : \mathbb{R}^d \to \mathbb{R}^d$ transports μ to ν in case $\nu(B) = \mu(\mathbf{t}^{-1}(B))$. Write $\mathbf{t}\#\mu = \nu$.



- $W_2(\mu, \nu) = \inf \left\{ \left(\int |\mathbf{t}(x) - x|^2 d\mu \right)^{1/2} : \mathbf{t}\#\mu = \nu \right\}$

Brenier-McCann: the inf is uniquely attained by the optimal transport map $\mathbf{t}_\mu^\nu$.

Geodesics and Convexity

Geodesics:

$\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν at time α,

$$W_2(\mu(\alpha), \mu(\beta)) = |\alpha - \beta| W_2(\mu, \nu) \ .$$

Geodesics and Convexity

Geodesics:

$\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν at time α,

$$W_2(\mu(\alpha), \mu(\beta)) = |\alpha - \beta| W_2(\mu, \nu) .$$

Convexity:

E is convex in case

$$E(\mu(\alpha)) \leq (1-\alpha)E(\mu(0)) + \alpha E(\mu(1)) .$$

Geodesics and Convexity

Geodesics:

$\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν at time α,

$$W_2(\mu(\alpha), \mu(\beta)) = |\alpha - \beta| W_2(\mu, \nu) \ .$$

Convexity:

E is convex in case

$$E(\mu(\alpha)) \leq (1-\alpha)E(\mu(0)) + \alpha E(\mu(1)) \ .$$

Geodesics and Convexity

Geodesics:

$\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν at time α,

$$W_2(\mu(\alpha), \mu(\beta)) = |\alpha - \beta| W_2(\mu, \nu) \; .$$

Convexity:

E is convex in case

$$E(\mu(\alpha)) \leq (1-\alpha)E(\mu(0)) + \alpha E(\mu(1)) \; .$$

Assumption: E is lower semicontinuous and convex.

The Wasserstein Metric's ``Gradient''

By a similar computation as in the L^2 case,

$$\left(\nabla_{W_2} E(\mu), \left. \frac{\partial \mu}{\partial t} \right|_{t=0} \right)_\mu = \lim_{t \to 0} \frac{E(\mu(t)) - E(\mu)}{t} \text{ for all } \frac{\partial \mu}{\partial t} ,$$

implies

$$\nabla_{W_2} E(\mu) = -\nabla \cdot \left(\mu \nabla \frac{\delta E}{\delta \mu} \right) .$$

The Wasserstein Metric's "Gradient"

By a similar computation as in the L^2 case,

$$\left(\nabla_{W_2} E(\mu),\, \frac{\partial \mu}{\partial t}\bigg|_{t=0}\right)_\mu = \lim_{t \to 0} \frac{E(\mu(t)) - E(\mu)}{t} \text{ for all } \frac{\partial \mu}{\partial t} ,$$

implies

$$\nabla_{W_2} E(\mu) = -\nabla \cdot \left(\mu \nabla \frac{\delta E}{\delta \mu}\right) .$$

Therefore,

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)) \iff \frac{\partial \mu(t)}{\partial t} = \nabla \cdot \left(\mu \nabla \frac{\delta E}{\delta \mu}\right) .$$

Examples of Wasserstein Gradient Flow

	PDE	Energy Functional
Porous Media / Fast Diffusion	$\frac{\partial}{\partial t}\mu = \Delta\mu^m$	$E(\mu) = \frac{1}{m-1}\int \rho(x)^m dx$
Fokker Planck	$\frac{\partial}{\partial t}\mu = \Delta\mu + \nabla\cdot(\mu\nabla V)$	$E(\mu) = \int \rho(x)\log\rho(x) + V(x)\rho(x)dx$
Aggregation	$\frac{\partial}{\partial t}u = \nabla\cdot(\mu\nabla K * \mu)$	$E(\mu) = \frac{1}{2}\int\int \rho(x)K(x-y)\rho(y)dxdy$

Examples of Wasserstein Gradient Flow

	PDE	Energy Functional
Porous Media / Fast Diffusion	$\frac{\partial}{\partial t}\mu = \Delta \mu^m$	$E(\mu) = \frac{1}{m-1}\int \rho(x)^m dx$
Fokker Planck	$\frac{\partial}{\partial t}\mu = \Delta \mu + \nabla \cdot (\mu \nabla V)$	$E(\mu) = \int \rho(x) \log \rho(x) + V(x)\rho(x) dx$
Aggregation	$\frac{\partial}{\partial t}u = \nabla \cdot (\mu \nabla K * \mu)$	$E(\mu) = \frac{1}{2}\int\int \rho(x) K(x-y) \rho(y) dx dy$

Why gradient flow?

- Free estimates, e.g. $W_2(\mu(t), \nu(t)) \le e^{-\lambda t} W_2(\mu(0), \nu(0))$
- Method to construct and approximate solutions (discrete gradient flow)

Plan

- Gradient flow
- Discrete gradient flow
- Euler-Lagrange equation
- Exponential formula

Discrete Gradient Flow: Euclidean Space

Gradient flow:

$$\frac{du(t)}{dt} = -\nabla E(u(t)), \quad u(0) = u \in \mathbb{R}^d$$

Discrete Gradient Flow: Euclidean Space

Gradient flow:

$$\frac{du(t)}{dt} = -\nabla E(u(t)), \quad u(0) = u \in \mathbb{R}^d$$

Implicit Euler method:

$$\frac{u_n - u_{n-1}}{\tau} = -\nabla E(u_n), \quad u_0 = u$$

Discrete Gradient Flow: Euclidean Space

Gradient flow:

$$\frac{du(t)}{dt} = -\nabla E(u(t)), \quad u(0) = u \in \mathbb{R}^d$$

Implicit Euler method:

$$\frac{u_n - u_{n-1}}{\tau} + \nabla E(u_n) = 0, \quad u_0 = u$$

Discrete Gradient Flow: Euclidean Space

Gradient flow:

$$\frac{du(t)}{dt} = -\nabla E(u(t)), \quad u(0) = u \in \mathbb{R}^d$$

Implicit Euler method:

$$\frac{u_n - u_{n-1}}{\tau} + \nabla E(u_n) = 0, \quad u_0 = u$$

Given u_{n-1}, compute u_n using that it is a critical point of

$$\Phi(v) = \frac{1}{2\tau}|v - u_{n-1}|^2 + E(v)\ .$$

Discrete Gradient Flow: Euclidean Space

Gradient flow:

$$\frac{du(t)}{dt} = -\nabla E(u(t)), \quad u(0) = u \in \mathbb{R}^d$$

Implicit Euler method:

$$\frac{u_n - u_{n-1}}{\tau} + \nabla E(u_n) = 0, \quad u_0 = u$$

Given u_{n-1}, compute u_n using that it is ~~a critical point~~ the unique minimizer of

$$\Phi(v) = \frac{1}{2\tau}|v - u_{n-1}|^2 + E(v) \ .$$

Discrete Gradient Flow: Euclidean Space

Gradient flow:

$$\frac{du(t)}{dt} = -\nabla E(u(t)), \quad u(0) = u \in \mathbb{R}^d$$

Implicit Euler method:

$$\frac{u_n - u_{n-1}}{\tau} + \nabla E(u_n) = 0, \quad u_0 = u$$

Given u_{n-1}, compute u_n using that it is ~~a critical point~~ the unique minimizer of

$$\Phi(v) = \frac{1}{2\tau}|v - u_{n-1}|^2 + E(v)\ .$$

Theorem (Exponential Formula)

Let $\tau = t/n$. Then $\lim_{n\to\infty} u_n = u(t)$.

Discrete Gradient Flow: Wasserstein Metric

Want to say... given μ_{n-1}, let μ_n be the unique minimizer of

$$\Phi(v) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \ .$$

Discrete Gradient Flow: Wasserstein Metric

Want to say... given μ_{n-1}, let μ_n be the unique minimizer of

$$\Phi(v) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \ .$$

Problem: $\nu \mapsto W_2^2(\nu, \mu_{n-1})$ is not convex, so Φ may not have a unique minimum.

Discrete Gradient Flow: Wasserstein Metric

Want to say... given μ_{n-1}, let μ_n be the unique minimizer of

$$\Phi(v) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \ .$$

Problem: $\nu \mapsto W_2^2(\nu, \mu_{n-1})$ is not convex, so Φ may not have a unique minimum.

Need additional assumptions on E:

- coercive
- *convex along generalized geodesics*

Discrete Gradient Flow: Wasserstein Metric

Want to say... given μ_{n-1}, let μ_n be the unique minimizer of

$$\Phi(v) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \ .$$

Problem: $\nu \mapsto W_2^2(\nu, \mu_{n-1})$ is not convex, so Φ may not have a unique minimum.

Need additional assumptions on E:

- coercive
- *convex along generalized geodesics*

Proposition (AGS)

For all $\tau > 0$, there exists a unique minimizer of $\Phi(\nu)$, so the discrete gradient flow is well defined.

Plan

- Gradient flow
- Discrete gradient flow
- Euler-Lagrange equation
- Exponential formula

Euler-Lagrange Equation

In the Euclidean case,

$$u_n = \operatorname*{argmin}_v \left\{ \frac{1}{2\tau} |v - u_{n-1}|^2 + E(v) \right\} \iff \frac{u_n - u_{n-1}}{\tau} = -\nabla E(u_n) \ .$$

Euler-Lagrange Equation

In the Euclidean case,

$$u_n = \operatorname*{argmin}_v \left\{ \frac{1}{2\tau} |v - u_{n-1}|^2 + E(v) \right\} \iff \frac{u_n - u_{n-1}}{\tau} = -\nabla E(u_n) \, .$$

In the Wasserstein case,

Proposition (AGS, C.)

$$\mu_n = \operatorname*{argmin}_\nu \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau} (\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \, .$$

Euler-Lagrange Equation

In the Euclidean case,

$$u_n = \operatorname*{argmin}_v \left\{ \frac{1}{2\tau} |v - u_{n-1}|^2 + E(v) \right\} \iff \frac{u_n - u_{n-1}}{\tau} = -\nabla E(u_n) \; .$$

In the Wasserstein case,

Proposition (AGS, C.)

$$\mu_n = \operatorname*{argmin}_\nu \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \; .$$

Key property of subdifferential: for E convex, $0 \in \partial E(\mu) \iff \mu$ minimizes E.

Sketch of Proof: Euler-Lagrange Equation

Proposition (AGS, C.)

$$\mu_n = \operatorname*{argmin}_{\nu} \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \; .$$

Sketch of Proof: Euler-Lagrange Equation

Proposition (AGS, C.)

$$\mu_n = \underset{\nu}{\operatorname{argmin}} \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \ .$$

Proof:

Let $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu)$.

Sketch of Proof: Euler-Lagrange Equation

Proposition (AGS, C.)

$$\mu_n = \underset{\nu}{\text{argmin}} \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \ .$$

Proof:

Let $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu)$.

$\implies$ [AGS, Otto]: minimality implies $\Phi(\mathbf{t}\#\mu_n) \geq \Phi(\mu_n)$; expand both sides.

Sketch of Proof: Euler-Lagrange Equation

Proposition (AGS, C.)

$$\mu_n = \operatorname*{argmin}_{\nu} \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \ .$$

Proof:

Let $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu)$.

$\Longrightarrow$ [AGS, Otto]: minimality implies $\Phi(\mathbf{t}\#\mu_n) \geq \Phi(\mu_n)$; expand both sides.

$\Longleftarrow$ want to say...

- $\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n)$
- hence $0 \in \partial\Phi(\mu_n)$
- hence by key property, μ_n minimizes Φ

Sketch of Proof: Euler-Lagrange Equation

Proposition (AGS, C.)

$$\mu_n = \operatorname*{argmin}_{\nu} \left\{ \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu) \right\} \iff \frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \,.$$

Proof:

Let $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu)$.

$\implies$ [AGS, Otto]: minimality implies $\Phi(\mathbf{t}\#\mu_n) \geq \Phi(\mu_n)$; expand both sides.

$\impliedby$ want to say...

- $\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n)$
- hence $0 \in \partial \Phi(\mu_n)$
- hence by key property, μ_n minimizes Φ

Problem: $\nu \mapsto W_2^2(\nu, \mu_{n-1})$ is not convex, so Φ may not be convex.

Sketch of Proof: Euler-Lagrange Equation

Solution: generalized geodesics and transport metrics

- [AGS] $\nu \mapsto W_2^2(\nu, \mu)$ is not convex (along all geodesics)
- [AGS] $\nu \mapsto W_2^2(\nu, \mu)$ is convex along generalized geodesics with base μ
- [C.] the generalized geodesics with base μ are not arbitrary curves: they are exactly the geodesics of the transport metric with base μ

Sketch of Proof: Euler-Lagrange Equation

Solution: generalized geodesics and transport metrics

- [AGS] $\nu \mapsto W_2^2(\nu, \mu)$ is not convex (along all geodesics)
- [AGS] $\nu \mapsto W_2^2(\nu, \mu)$ is convex along generalized geodesics with base μ
- [C.] the generalized geodesics with base μ are not arbitrary curves: they are exactly the geodesics of the transport metric with base μ

Wasserstein Metric: $W_2(\mu, \nu) = \left(\int |\mathbf{t}_\mu^\nu - \mathbf{id}|^2 d\mu\right)^{1/2}$

Transport Metric: $W_{2,\omega}(\mu, \nu) = \left(\int |\mathbf{t}_\omega^\mu - \mathbf{t}_\omega^\nu|^2 d\omega\right)^{1/2}$

- $\nu \mapsto W_{2,\omega}^2(\nu, \mu)$ is convex
- $W_2(\mu, \nu) \leq W_{2,\omega}(\mu, \nu)$

Plan

- Gradient flow
- Discrete gradient flow
- Euler-Lagrange equation
- Exponential formula

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$*. Then* $\lim_{n\to\infty} \mu_n = \mu(t)$.

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$. *Then* $\lim_{n\to\infty} \mu_n = \mu(t)$.

- the limit exists
- the limit is a solution to the gradient flow

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$. *Then* $\lim_{n\to\infty} \mu_n = \mu(t)$.

- the limit exists
- the limit is a solution to the gradient flow

Sketch of Proof, a la Crandall and Liggett:

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$. Then $\lim_{n\to\infty} \mu_n = \mu(t)$*.*

- the limit exists
- the limit is a solution to the gradient flow

Sketch of Proof, a la Crandall and Liggett:

Let J_τ be the function $J_\tau u = \text{argmin}_v \left\{ \frac{1}{2\tau} |v-u|^2 + E(v) \right\} \implies J_\tau^n u_0 = u_n$.

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$. Then $\lim_{n\to\infty} \mu_n = \mu(t)$.

- the limit exists
- the limit is a solution to the gradient flow

Sketch of Proof, a la Crandall and Liggett:

Let J_τ be the function $J_\tau u = \operatorname{argmin}_v \left\{ \frac{1}{2\tau}|v-u|^2 + E(v) \right\} \implies J_\tau^n u_0 = u_n$.

1. **Contraction inequality**

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$. *Then* $\lim_{n\to\infty} \mu_n = \mu(t)$.

- the limit exists
- the limit is a solution to the gradient flow

Sketch of Proof, a la Crandall and Liggett:

Let J_τ be the function $J_\tau u = \operatorname{argmin}_v \left\{ \frac{1}{2\tau}|v-u|^2 + E(v) \right\} \implies J_\tau^n u_0 = u_n$.

1 **Contraction inequality**

Banach space: $\|J_\tau u - J_\tau v\| \leq \|u - v\|$

Exponential Formula

Theorem (AGS)

Let $\tau = t/n$. *Then* $\lim_{n\to\infty} \mu_n = \mu(t)$.

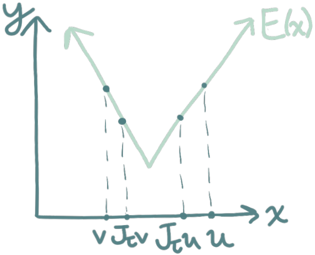
- the limit exists
- the limit is a solution to the gradient flow

Sketch of Proof, a la Crandall and Liggett:

Let J_τ be the function $J_\tau u = \text{argmin}_v \left\{ \frac{1}{2\tau}|v-u|^2 + E(v) \right\} \implies J_\tau^n u_0 = u_n$.

1. **Contraction inequality**

Banach space: $\|J_\tau u - J_\tau v\| \leq \|u - v\|$



Exponential Formula

Theorem (AGS)

Let $\tau = t/n$*. Then* $\lim_{n\to\infty} \mu_n = \mu(t)$.

- the limit exists
- the limit is a solution to the gradient flow

Sketch of Proof, a la Crandall and Liggett:

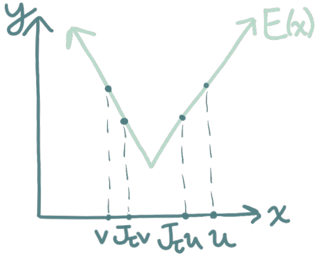
Let J_τ be the function $J_\tau u = \operatorname{argmin}_v \left\{ \frac{1}{2\tau} |v-u|^2 + E(v) \right\} \implies J_\tau^n u_0 = u_n$.

1 **Contraction inequality**

Banach space: $\|J_\tau u - J_\tau v\| \leq \|u - v\|$

Theorem (Carlen, C.)

$W_2^2(J_\tau \mu, J_\tau \nu) \leq W_2^2(\mu, \nu) + \mathcal{O}(\tau^2)$



Exponential Formula

② **Large vs small time steps,** $0 < h \le \tau$

Exponential Formula

② **Large vs small time steps,** $0 < h \leq \tau$

Banach space: $J_\tau u = J_h \left[\frac{\tau - h}{\tau} J_\tau u + \frac{h}{\tau} u \right]$

Exponential Formula

2 Large vs small time steps, $0 < h \leq \tau$

Banach space: $J_\tau u = J_h \left[\frac{\tau - h}{\tau} J_\tau u + \frac{h}{\tau} u\right]$

Lemma (Jost, Mayer, C.)

$J_\tau \mu = J_h \left[\left(\frac{\tau-h}{\tau} t_\mu^{J_\tau \mu} + \frac{h}{\tau} \mathbf{id}\right) \# \mu\right]$

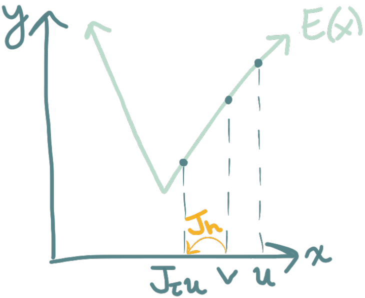
Exponential Formula

② **Large vs small time steps,** $0 < h \leq \tau$

Banach space: $J_\tau u = J_h \left[\frac{\tau - h}{\tau} J_\tau u + \frac{h}{\tau} u\right]$

Lemma (Jost, Mayer, C.)

$J_\tau \mu = J_h \left[\left(\frac{\tau-h}{\tau} t_\mu^{J_\tau \mu} + \frac{h}{\tau}\mathbf{id}\right)\#\mu\right]$



Exponential Formula

3 **Recursive inequality**:

Exponential Formula

③ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

Exponential Formula

❸ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$W_2^2(\mu_n, \mu_m) = W_2^2\left(J_h \overbrace{\left[\left(\frac{\tau - h}{\tau} t_{\mu_{n-1}}^{\mu_n} + \frac{h}{\tau} \textbf{id} \right) \# \mu_{n-1} \right]}^{\nu}, J_h \mu_{m-1} \right)$$

Exponential Formula

❸ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$W_2^2(\mu_n, \mu_m) = W_2^2\left(J_h \overbrace{\left[\left(\frac{\tau - h}{\tau} t_{\mu_{n-1}}^{\mu_n} + \frac{h}{\tau}\textbf{id} \right) \# \mu_{n-1} \right]}^{\nu}, J_h \mu_{m-1} \right)$$

$$\leq W_2^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2)$$

Exponential Formula

3 **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$
\begin{aligned}
W_2^2(\mu_n, \mu_m) &= W_2^2\left(J_h \overbrace{\left[\left(\frac{\tau - h}{\tau} t_{\mu_{n-1}}^{\mu_n} + \frac{h}{\tau}\textbf{id} \right) \#\mu_{n-1} \right]}^{\nu}, J_h \mu_{m-1} \right) \\
&\leq W_2^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\
&\leq W_{2,\mu_{n-1}}^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2)
\end{aligned}
$$

Exponential Formula

③ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$
\begin{aligned}
W_2^2(\mu_n, \mu_m) &= W_2^2\left(J_h \overbrace{\left[\left(\frac{\tau - h}{\tau} t_{\mu_{n-1}}^{\mu_n} + \frac{h}{\tau}\textbf{id} \right) \# \mu_{n-1} \right]}^{\nu}, J_h \mu_{m-1} \right) \\
&\leq W_2^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\
&\leq W_{2,\mu_{n-1}}^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\
&\leq \frac{\tau - h}{\tau} W_{2,\mu_{n-1}}^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_{2,\mu_{n-1}}^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2)
\end{aligned}
$$

Exponential Formula

❸ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$\begin{aligned}
W_2^2(\mu_n, \mu_m) &= W_2^2\left(J_h \overbrace{\left[\left(\frac{\tau - h}{\tau} t_{\mu_{n-1}}^{\mu_n} + \frac{h}{\tau} \textbf{id} \right) \# \mu_{n-1} \right]}^{\nu}, J_h \mu_{m-1} \right) \\
&\leq W_2^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\
&\leq W_{2,\mu_{n-1}}^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\
&\leq \frac{\tau - h}{\tau} W_{2,\mu_{n-1}}^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_{2,\mu_{n-1}}^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2) \\
&\leq \frac{\tau - h}{\tau} W_2^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_2^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2)
\end{aligned}$$

Exponential Formula

③ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$\begin{aligned} W_2^2(\mu_n, \mu_m) &= W_2^2\left(J_h \overbrace{\left[\left(\frac{\tau - h}{\tau} t_{\mu_{n-1}}^{\mu_n} + \frac{h}{\tau}\mathbf{id}\right)\#\mu_{n-1}\right]}^{\nu}, J_h \mu_{m-1} \right) \\ &\le W_2^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\ &\le W_{2,\mu_{n-1}}^2(\nu, \mu_{m-1}) + \mathcal{O}(h^2) \\ &\le \frac{\tau - h}{\tau} W_{2,\mu_{n-1}}^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_{2,\mu_{n-1}}^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2) \\ &\le \frac{\tau - h}{\tau} W_2^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_2^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2) \end{aligned}$$

$$W_2^2(\mu_n, \mu_m) \le \frac{\tau - h}{\tau} W_2^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_2^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2)$$

Exponential Formula

③ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

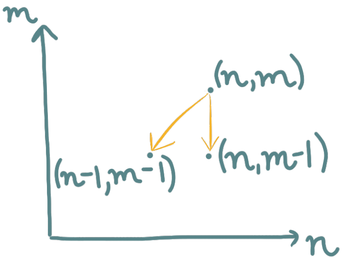
$$W_2^2(\mu_n, \mu_m) \leq \frac{\tau - h}{\tau} W_2^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_2^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2)$$

Exponential Formula

③ **Recursive inequality**:

Let $\mu_n = J_\tau^n \mu$ and $\mu_m = J_h^m \mu$.

$$W_2^2(\mu_n, \mu_m) \le \frac{\tau - h}{\tau} W_2^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_2^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2)$$



Exponential Formula

Iterating

$$W_2^2(\mu_n, \mu_m) \leq \frac{\tau - h}{\tau} W_2^2(\mu_n, \mu_{m-1}) + \frac{h}{\tau} W_2^2(\mu_{n-1}, \mu_{m-1}) + \mathcal{O}(h^2)$$

with $\tau = t/n$ and $h = t/m$ for $n \leq m$ gives

$$W_2(\mu_n, \mu_m) \leq \mathcal{O}(\frac{1}{\sqrt{n}}) \xrightarrow{n,m\to\infty} 0 \,.$$

Therefore, the limit exists. $\square$

Thank you!

Backup

Wasserstein Gradient Flow

$$\frac{\partial \mu(t)}{\partial t} = -\nabla_{W_2} E(\mu(t)), \quad \mu(0) = \mu$$

Wasserstein Metric as ``Riemannian Manifold''*

The Wasserstein metric is induced by this inner product (Benamou-Brenier):

$$W_2(\mu_0, \mu_1) =$$
$$\inf \left\{ \int_0^1 \|\nabla \psi(t)\|_{\mu(t)} dt : \mu(0) = \mu_0,\ \mu(1) = \mu_1,\ \frac{\partial \mu}{\partial t} + \nabla \cdot (\nabla \psi \mu) = 0 \right\} .$$

The Wasserstein Metric's ``Inner Product''* [Otto]

The Wasserstein Metric's ``Inner Product''* [Otto]

*DISCLAIMER: "given sufficient regularity," "in the limit", ...

The Wasserstein Metric's ``Inner Product''* [Otto]

*DISCLAIMER: "given sufficient regularity," "in the limit", ...

Given $\mu(t)$, there exists a velocity field $v(x,t) = \nabla\psi(x,t)$ so that

$$\frac{\partial \mu}{\partial t} + \nabla \cdot (\nabla \psi \mu) = 0 \ .$$

The Wasserstein Metric's ``Inner Product''* [Otto]

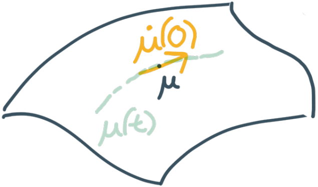
*DISCLAIMER: "given sufficient regularity," "in the limit", ...

Given $\mu(t)$, there exists a velocity field $v(x,t) = \nabla\psi(x,t)$ so that

$$\frac{\partial \mu}{\partial t} + \nabla \cdot (\nabla \psi \mu) = 0 \; .$$

The tangent space at a measure μ is

$$\left\{ \left. \frac{\partial \mu}{\partial t} \right|_{t=0} : \mu(0) = \mu \right\} = \left\{ \nabla \psi : \psi \in C_c^\infty(\mathbb{R}^d) \right\} \; .$$



The Wasserstein Metric's ``Inner Product''* [Otto]

*DISCLAIMER: "given sufficient regularity," "in the limit", ...

Given $\mu(t)$, there exists a velocity field $v(x,t) = \nabla\psi(x,t)$ so that

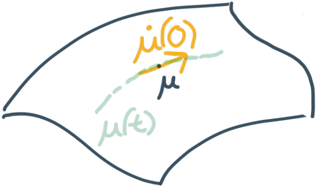
$$\frac{\partial \mu}{\partial t} + \nabla \cdot (\nabla \psi \mu) = 0 \ .$$

The tangent space at a measure μ is

$$\left\{ \left. \frac{\partial \mu}{\partial t} \right|_{t=0} : \mu(0) = \mu \right\} = \left\{ \nabla \psi : \psi \in C_c^\infty(\mathbb{R}^d) \right\} \ .$$

The inner product is

$$\left(\frac{\partial \mu}{\partial t}, \frac{\tilde{\partial} \mu}{\partial t} \right)_\mu := \int \nabla \psi(x) \cdot \nabla \tilde{\psi}(x) d\mu \ .$$



Wasserstein Subdifferential

Wasserstein subdifferential of convex function:

- $\boldsymbol{\xi} \in \partial E(\mu)$ in case $E(\nu) - E(\mu) \geq \int \langle \boldsymbol{\xi}, \mathbf{t}_\mu^\nu - \mathbf{id} \rangle d\mu$ for all ν
- $\boldsymbol{\xi} \in \partial_s E(\mu)$ in case $E(\nu) - E(\mu) \geq \int \langle \boldsymbol{\xi}, \mathbf{t} - \mathbf{id} \rangle d\mu$ for all ν and all $\mathbf{t}\#\mu = \nu$.

Generalized Geodesics

- $\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν
- $\mu(\alpha) = (\alpha \mathbf{t}_\omega^\mu + (1-\alpha)\mathbf{t}_\omega^\nu)\#\omega$ is the gen. geodesic from μ to ν with base ω

Generalized Geodesics

- $\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν
- $\mu(\alpha) = (\alpha \mathbf{t}_\omega^\mu + (1-\alpha)\mathbf{t}_\omega^\nu)\#\omega$ is the gen. geodesic from μ to ν with base ω

Proposition (AGS)

$\nu \mapsto W_2^2(\nu, \mu)$ is convex along gen. geodesics with base μ.

Generalized Geodesics

- $\mu(\alpha) = (\alpha \mathbf{t}_\mu^\nu + (1-\alpha)\mathbf{id})\#\mu$ is the geodesic from μ to ν
- $\mu(\alpha) = (\alpha \mathbf{t}_\omega^\mu + (1-\alpha)\mathbf{t}_\omega^\nu)\#\omega$ is the gen. geodesic from μ to ν with base ω

Proposition (AGS)

$\nu \mapsto W_2^2(\nu,\mu)$ is convex along gen. geodesics with base μ.

Thus, E convex along gen. geodesics $\implies$
$\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu)$ convex along gen. geodesics with base μ_{n-1}.

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega, \nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. *$W_{2,\mu}$ is a metric*

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega, \nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. $W_{2,\mu}$ *is a metric*
2. $W_2(\omega, \nu) \leq W_{2,\mu}(\omega, \nu)$ *with equality if* $\mu = \omega$ *or* $\mu = \nu$

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. $W_{2,\mu}$ *is a metric*
2. $W_2(\omega,\nu) \leq W_{2,\mu}(\omega,\nu)$ *with equality if* $\mu = \omega$ *or* $\mu = \nu$
3. *the geodesics of* $W_{2,\mu}$ *are the gen. geodesics with base* μ

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. *$W_{2,\mu}$ is a metric*
2. *$W_2(\omega,\nu) \le W_{2,\mu}(\omega,\nu)$ with equality if $\mu = \omega$ or $\mu = \nu$*
3. *the geodesics of $W_{2,\mu}$ are the gen. geodesics with base μ*
4. *$\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ is convex*

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. *$W_{2,\mu}$ is a metric*
2. *$W_2(\omega,\nu) \le W_{2,\mu}(\omega,\nu)$ with equality if $\mu = \omega$ or $\mu = \nu$*
3. *the geodesics of $W_{2,\mu}$ are the gen. geodesics with base μ*
4. *$\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ is convex*
5. $\boldsymbol{\xi} \in \partial_s E(\nu) \implies \boldsymbol{\xi} \circ \mathbf{t}_\mu^\nu \in \partial_{2,\mu} E(\nu)$

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. $W_{2,\mu}$ *is a metric*
2. $W_2(\omega,\nu) \le W_{2,\mu}(\omega,\nu)$ *with equality if* $\mu = \omega$ *or* $\mu = \nu$
3. *the geodesics of* $W_{2,\mu}$ *are the gen. geodesics with base* μ
4. $\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ *is convex*
5. $\boldsymbol{\xi} \in \partial_s E(\nu) \implies \boldsymbol{\xi} \circ \mathbf{t}_\mu^\nu \in \partial_{2,\mu} E(\nu)$

Proof of Euler-Lagrange equation:

$\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \implies \mu_n = \mathrm{argmin}_\nu \left\{\frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)\right\}$.

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. *$W_{2,\mu}$ is a metric*
2. *$W_2(\omega,\nu) \leq W_{2,\mu}(\omega,\nu)$ with equality if $\mu = \omega$ or $\mu = \nu$*
3. *the geodesics of $W_{2,\mu}$ are the gen. geodesics with base μ*
4. *$\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ is convex*
5. $\boldsymbol{\xi} \in \partial_s E(\nu) \implies \boldsymbol{\xi} \circ \mathbf{t}_\mu^\nu \in \partial_{2,\mu} E(\nu)$

Proof of Euler-Lagrange equation:

$\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \implies \mu_n = \mathrm{argmin}_\nu \left\{\frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)\right\}$.

- E convex along gen. geodesics $\implies$ convex in $W_{2,\mu_{n-1}}$

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. $W_{2,\mu}$ *is a metric*
2. $W_2(\omega,\nu) \leq W_{2,\mu}(\omega,\nu)$ *with equality if* $\mu = \omega$ *or* $\mu = \nu$
3. *the geodesics of* $W_{2,\mu}$ *are the gen. geodesics with base* μ
4. $\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ *is convex*
5. $\boldsymbol{\xi} \in \partial_s E(\nu) \implies \boldsymbol{\xi} \circ \mathbf{t}_\mu^\nu \in \partial_{2,\mu} E(\nu)$

Proof of Euler-Lagrange equation:

$\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \implies \mu_n = \text{argmin}_\nu \left\{\frac{1}{2\tau} W_2^2(\nu, \mu_{n-1}) + E(\nu)\right\}$.

- E convex along gen. geodesics $\implies$ convex in $W_{2,\mu_{n-1}}$
- $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)$ convex in $W_{2,\mu_{n-1}}$

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. $W_{2,\mu}$ *is a metric*
2. $W_2(\omega,\nu) \leq W_{2,\mu}(\omega,\nu)$ *with equality if* $\mu = \omega$ *or* $\mu = \nu$
3. *the geodesics of* $W_{2,\mu}$ *are the gen. geodesics with base* μ
4. $\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ *is convex*
5. $\boldsymbol{\xi} \in \partial_s E(\nu) \implies \boldsymbol{\xi} \circ \mathbf{t}_\mu^\nu \in \partial_{2,\mu} E(\nu)$

Proof of Euler-Lagrange equation:

$\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \implies \mu_n = \text{argmin}_\nu \left\{\frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)\right\}$.

- E convex along gen. geodesics $\implies$ convex in $W_{2,\mu_{n-1}}$
- $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)$ convex in $W_{2,\mu_{n-1}}$
- Since $\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n)$, a computation shows $0 \in \partial_{2,\mu_{n-1}} \Phi(\mu_n)$

Transport Metrics

The transport metric with base μ is $W_{2,\mu}(\omega,\nu) := \left(\int |\mathbf{t}_\mu^\omega - \mathbf{t}_\mu^\nu|^2 d\mu\right)^{1/2}$.

Proposition (C.)

1. *$W_{2,\mu}$ is a metric*
2. *$W_2(\omega,\nu) \leq W_{2,\mu}(\omega,\nu)$ with equality if $\mu = \omega$ or $\mu = \nu$*
3. *the geodesics of $W_{2,\mu}$ are the gen. geodesics with base μ*
4. *$\omega \mapsto W_{2,\mu}^2(\omega,\nu)$ is convex*
5. $\boldsymbol{\xi} \in \partial_s E(\nu) \implies \boldsymbol{\xi} \circ \mathbf{t}_\mu^\nu \in \partial_{2,\mu} E(\nu)$

Proof of Euler-Lagrange equation:

$\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n) \implies \mu_n = \mathrm{argmin}_\nu \left\{\frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)\right\}$.

- E convex along gen. geodesics $\implies$ convex in $W_{2,\mu_{n-1}}$
- $\Phi(\nu) = \frac{1}{2\tau} W_2^2(\nu,\mu_{n-1}) + E(\nu)$ convex in $W_{2,\mu_{n-1}}$
- Since $\frac{1}{\tau}(\mathbf{t}_{\mu_n}^{\mu_{n-1}} - \mathbf{id}) \in \partial_s E(\mu_n)$, a computation shows $0 \in \partial_{2,\mu_{n-1}} \Phi(\mu_n)$
- Therefore, μ_n minimizes Φ.