Optimal Transportation: Fluid Dynamics View

David Gu

Computer Science Department
Stony Brook University

gu@cs.stonybrook.edu

September 7, 2022

Overview

There are three views of optimal transportation theory:

1. Duality view
2. Fluid dynamics view
3. Differential geometric view

Different views give different insights and induce different computational methods; but all three theories are coherent and consistent.

Optimal Transportation Map

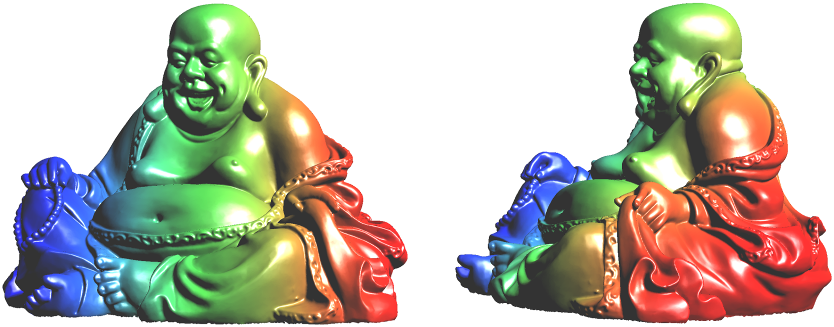

Figure: Buddha surface.

Optimal Transportation Map

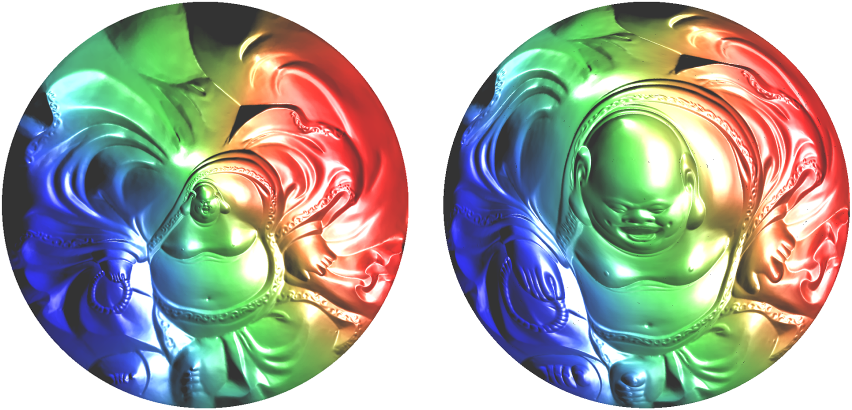

Figure: Optimal transportation map.

Optimal Transportation Map

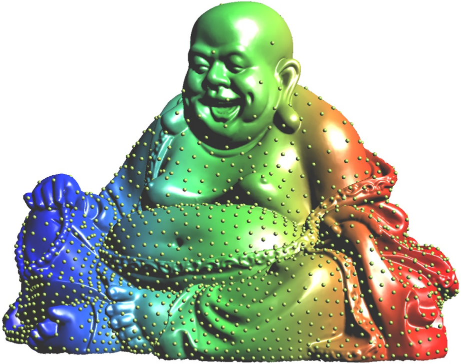

Figure: Brenier potential.

Optimal Transportation Map

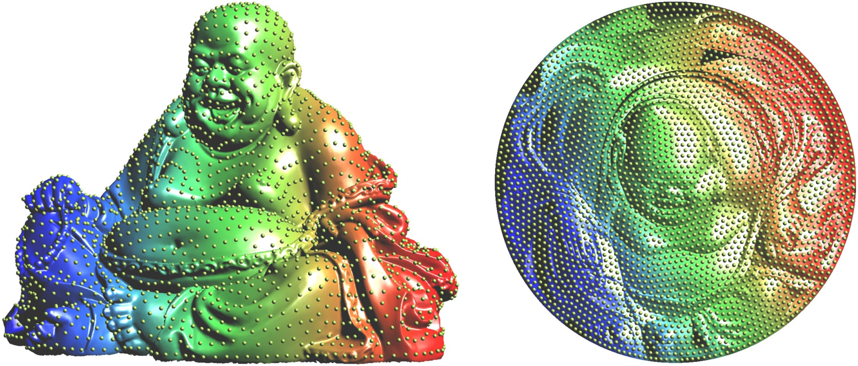

Figure: Brenier potential.

Optimal Transportation Map

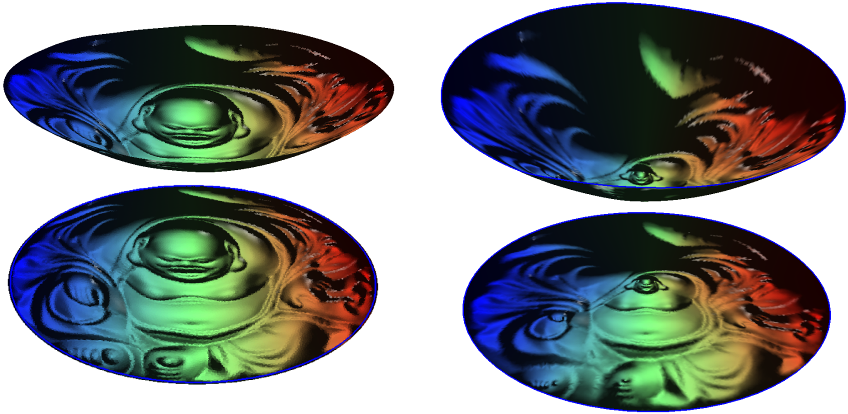

Figure: Brenier potential.

Euler Equation

The incompressible Euler Equation

We shall now study the dynamics of fluid flows and consider changes in motion due to forces acting on a fluid. We derive an evolution equation for the fluid momentum by considering forces acting on a small domain of fluid, of volume V and surface S, containing many fluid particles.

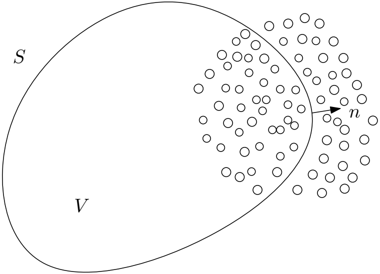

Figure: Forces acting on fluid volume.

Material Derivative

Consider a scalar quantity $\varphi = \varphi(\mathbf{x}, t)$, where t is time and $\mathbf{x}$ is position. Here φ may be some physical variable such as temperature or chemical concentration. The physical quantity, whose scalar quantity is φ, exists in a continuum, and whose macroscopic velocity is represented by the vector field $\mathbf{u}(\mathbf{x}, t)$.

The (total) derivative with respect to time of φ is expanded using the multivariate chain rule:

$$\frac{d}{dt}\varphi(\mathbf{x}, t) = \frac{\partial \varphi}{\partial t} + \dot{\mathbf{x}} \cdot \nabla \varphi.$$

It is apparent that this derivative is dependent on the vector $\dot{\mathbf{x}} = d\mathbf{x}/dt$, which describes a chosen path $\mathbf{x}(t)$ in space.

Material Derivative

The material derivative finally is obtained when the path $\mathbf{x}(t)$ is chosen to have a velocity equal to the fluid velocity $\dot{\mathbf{x}} = \mathbf{u}$. That is, the path follows the fluid current described by the fluid's velocity field $\mathbf{u}$. So, the *material derivative* of the scalar φ is

$$\frac{\mathrm{D}\varphi}{\mathrm{D}t} = \frac{\partial \varphi}{\partial t} + \mathbf{u} \cdot \nabla \varphi.$$

Similarly, the material derivative of $\mathbf{u}$ is

$$\boxed{\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u}}$$

Forces acting on a fluid

The forces acting on the fluid can be divided into two types. *Body forces*, such as *gravity*, act on all the particles throughout V,

$$F_V = \int_V \rho \mathbf{g} dV.$$

Surface forces are caused by interactions at the surface S, such as fluid *pressure*. Collisions between fluid molecules on either sides of the surface S produce a flux of momentum across the boundary, in the direction of the normal $\mathbf{n}$. The force exerted on the fluid into V by the fluid on the other side of S, by convention, written as

$$F_s = \int_S -p\mathbf{n} dS.$$

Euler's Equation

Newton's second law of motion tells that the sum of the forces acting on the volume of fluid V is equal to the rate of change of its momentum. Since $D\mathbf{v}/Dt$ is the acceleration of the fluid particles, or fluid elements, within V, one has

$$\int_V \rho \frac{D\mathbf{v}}{dt} dV = \int_S -p\mathbf{n} dS + \int_V \rho \mathbf{g} dV.$$

We now apply the divergence theorem,

$$\int_V \rho \frac{D\mathbf{v}}{Dt} dV = \int_V (-\nabla p + \rho \mathbf{g}) dV,$$

and notice that both integrands must be identical, since V is arbitrary. So the eveloution of fluid momentum is governed by *Euler's equation*

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = -\nabla p + \rho \mathbf{g}. \tag{1}$$

Euler Equation

The Euler equation is one of the most basic equations in fluid mechanics. In its simplest version it models an incompressible, inviscid fluid in a bounded, smooth open set $\Omega \subset \mathbb{R}^n$. The unknown is the velocity field of the fluid,

$$v(t,x) : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n$$

and the incompressible Euler equation is obtained by Newton's law, the material derivative of v is given by the gradient of the pressure, $Dv/Dt = -\nabla p$, from (1), we obtain

$$\boxed{\left\{\begin{array}{lcl} \frac{\partial v}{\partial t} + v \cdot \nabla v & = & -\nabla p \\ \nabla \cdot v & = & 0. \end{array}\right.} \tag{2}$$

Because $v(t,x)$ is divergence free, ρ is constant.

Energy Conservation

Lemma (Energy conservation)

Let v be a smooth solution of (2). Then, the total kinetic energy

$$\int_\Omega |v(t,x)|^2 dx = \|v(t,\cdot)\|^2_{L^2(\Omega;\mathbb{R}^n)}$$

is constant.

Proof.

Here, we give a sketch of proof.

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |v|^2 = \int_\Omega v \cdot \frac{\partial v}{\partial t} = -\int_\Omega v \cdot (v \cdot \nabla v) - \int_\Omega v \cdot \nabla p$$

But on one hand, because v is tangent to the boundary of Ω, we have

$$\int_\Omega v \cdot \nabla p = -\int_\Omega (\nabla \cdot v) p,$$

Energy Conservation

continued.

since $\nabla \cdot v = 0$, the above equals to 0. And on the other hand,

$$\int_\Omega v \cdot (v \cdot \nabla v) = \sum_{1 \le i,j \le n} \int_\Omega v_i v_j \partial_j v_i = \frac{1}{2} \sum_{ij} \int_\Omega v_j \partial_j (v_i)^2 = \frac{1}{2} \int_\Omega v \cdot \nabla |v|^2$$
$$= -\frac{1}{2} \int_\Omega (\nabla \cdot v) |v|^2 = 0.$$

Lagrange Formulation

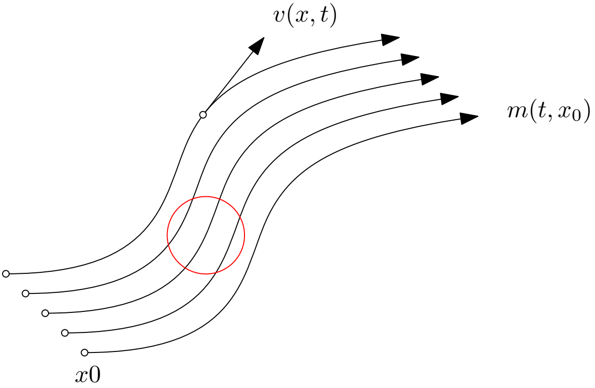

Figure: Lagrangian formulation.

Lagrangian Formulation

The point of view used in the last subsection, in which the unknown was a time-dependent velocity field, is called the *Eulerian* formulation. There is an alternative way of looking at fluid mechanics: the *Lagrangian* point of view, which focuses on the trajectories of particles.

- In an Eulerian description, one stares at a given, fixed point of space x, and measure the velocity $v(t, x)$ of fluid particles going through this point at time t.
- In a Lagrangian description, one puts a label on each particle, and then studies the trajectory of each labelled particle. For instance, assuming that we label particles according to their initial position x_0, we denote by

$$x = m(t, x_0)$$

the position at time t of a particle that was located at position x_0 at time 0. It is usually assumed that for each time t, the map $x_0 \mapsto m(t, x_0)$, defined on Ω, is one-to-one.

Eulerian and Lagrangian Views

To switch between these two descriptions, it suffices to use the identities

$$\left\{\begin{array}{rcl} v(t, m(t, x_0)) & = & \frac{d}{dt} m(t, x_0), \\ m(0, x_0) & = & x_0. \end{array}\right. \tag{3}$$

It is important to keep in mind the Eulerian expression of the *Lagrangian acceleration*: by differentiating (3) with respect to time, one finds that

$$\frac{d^2}{dt^2} m(t, x_0) = \left[\frac{\partial v}{\partial t} + (v \cdot \nabla v)\right](t, m(t, x_0)), \tag{4}$$

which explains the occurence of the convective derivative $v \cdot \nabla v$ in the Euler equation.

Diffeomorphism

It seems natural to search for $(m(t,\cdot))_{t\geq 0}$ as a family of *diffeomorphisms* from Ω to Ω:

- The mapping m is injective, because different trajectories won't intersect at the same time t;
- the mapping is surjective, otherwise there would be some vaccum created inside the domain, contradicting the fact that the fluid has constant density (incompressibility).

Diffeomorphism

The imcompressibility constraint can be recast in terms of m: if v is regular enough, then

$$\nabla \cdot v = 0 \iff \det\left(\frac{\partial m}{\partial x_0}\right) \equiv 1. \tag{5}$$

Indeed, the identity on the right-hand side of (5) is obviously satisfied at time 0, since $m(0,\cdot)$ is the identity map; then (5) is a consequence of the identity

$$\frac{\partial}{\partial t}\log\det\left[\frac{\partial m}{\partial x_0}\right] = (\nabla \cdot v)(t, m(t, x_0)). \tag{6}$$

The formula can be proved using the derivative of the determinant of a matrix,

$$\frac{d}{dt}\det(A(t)) = \det(A(t))\, Tr\left(A^{-1}\frac{d}{dt}A\right).$$

Diffeomorphism

In Lagrangian formulation, the Euler equation becomes an evolution equation for a map $t \mapsto m(t, \cdot)$, with values in the group $G(\Omega)$ of diffeomorphisms $\Omega \to \Omega$ with unit determininant. To recall this, we use the letter g for the trajectory map m. In particular, g is *measure-preserving*: it pushes Lebesgue measure forward to itself. The physical interpretation is that the volume of a set of particles is kept constant under time-evolution, which is precisely the *incompressibility*. Thus, we rewrite (3) as

$$v(t, g(t, x_0)) = \frac{d}{dt} g(t, x_0), \quad \text{or} \quad v(t.x) = \frac{\partial g}{\partial t} \circ g^{-1}(t, x). \tag{7}$$

By (4), the Euler equation (2) translates into an equation on the trajectory field $t \mapsto g(t, \cdot)$ of $\mathbb{R}_+$ into $G(\Omega)$,

$$\frac{d^2}{dt^2} g(t, x_0) = -\nabla p(t, g(t, x_0)). \tag{8}$$

Arnold's Interpretation

Theorem (Arnold)

The Euler equation is the equation of geodesics on $G(\Omega)$, endowed with the Riemannian structure inherited from the Euclidean space $L^2(\Omega;\mathbb{R}^n)$.

Recall that a geodesic on a Riemannian manifold M is a path $\gamma(t)$ which minimizes the distance

$$\left[\int_{t_1}^{t_2} |\dot{g}(t)|^2 dt\right]^{\frac{1}{2}} \tag{9}$$

among all curves $g:[t_1,t_2]\to M$ constrained by the boundary conditions $g(t_1)=\gamma(t_1)$, $g(t_2)=\gamma(t_2)$, and this minimization property should hold true whenever t_2 is close enough to t_1. It is equivalent to stating that the acceleration of the curve γ, viewed from the tangent space to the manifold, vanishes identically.

Arnold's Interpretation

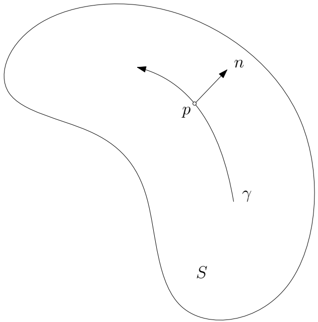

Figure: Geodesic: curve normal coincide with the surface normal.

In Arnold's interpretation, we consider the Riemannian structure on $G(\Omega)$ inherited from $L^2(\Omega)$, and this simply means that the acceleration d^2g/dt^2 should be orthogonal to the tangent space $T_{g(t)}G(\Omega)$ in $L^2(\Omega; \mathbb{R}^n)$. Recall from the preceding discussion that a path $g(t)$, starting from $g_0 \in G$, stays in G if and only if $\partial g/\partial t$ is tangent to the boundary, and $\nabla \cdot v(t,x) = 0$, by (7)

$$\nabla \cdot \left[\frac{\partial g}{\partial t} \circ g^{-1} \right] = 0.$$

Thus, tangent vector in $T_g G$ are all vector fields h such that $\nabla \cdot (h \circ g^{-1}) = 0$, or equivalently $h = w_0 \circ g$, where w_0 lies in D_0, the space of divergence-free vector fields.

Arnold's Interpretation

Using the fact that g is measure-preserving diffeomorphism, one immediately checks that $(T_g G)^{\perp}$ is the space of all vector fields $q_0 \circ g$, where $g_0 \in D_0^{\perp}$, and $D_0^{\perp}$ is the orthogonal subspace to D_0 in L^2. Now, it is an easy consequence of the Helmoltz decomposition that under reasonable regularity condition on Ω,

$$D_0^{\perp} = \{-\nabla p,\ p : \Omega \to \mathbb{R}\}.$$

So the equation for geodesics becomes

$$\frac{d^2}{dt^2} g(t) = -\nabla p(t, g(t)).$$

This is exactly the incompressible Euler equation (8) in Lagrangian formulation. This gives an interpretation of the pressure field.

Arnold's Interpretation

If $g(t, x_0)$ is a smooth solution of the Euler equation, with a pressure field p bounded in $C^2(\Omega)$, uniformly in time, then there exists $\varepsilon > 0$ such that for $|t_1 - t_2| < \varepsilon$,

$$\int_{t_1}^{t_2} \left(\int_\Omega \left| \frac{\partial g}{\partial t}(t, x_0) \right|^2 dx_0 \right) dt \leq \int_{t_1}^{t_2} \left(\int_\Omega \left| \frac{\partial \gamma}{\partial t}(t, x_0) \right|^2 dx_0 \right) dt$$

for any other trajectory mapping γ with $\gamma(t_1) = g(t_1)$ and $\gamma(t_2) = g(t_2)$. If Ω is convex, one can choose

$$\varepsilon = \frac{\pi}{\sqrt{\|D^2 p\|_{L^\infty}}}.$$

Fluid Dynamics View

Fluid View

Consider a flow field of special gas. At each time $t \in [0,1]$, at point $x \in \Omega$, the density of the gas is $\rho(x,t)$. For Lagrangian point of view, the trajectory of each particle (molecule) is a curve, denoted as $\gamma_x(t)$, with initial position and velocity

$$\gamma_x(0) = x, \quad \gamma_x'(t) = \mathbf{v}(\gamma_x(t)).$$

The density function is $\rho(x,t)$, by mass conservation law, we get the continuity equation

$$\boxed{\partial_t \rho_t + \nabla \cdot (\rho_t \mathbf{v}_t) = 0}.$$

Diffeomorphism

Suppose the trajectories of different particles intersect at some time t, then globally there will be shock waves in the flow field. If there is no shocks, then at each time $t \in [0, 1]$, the initial position x of each particle is mapped to the current position $\gamma_x(t)$, this gives a global diffeomorphism

$$g_t := g(\cdot, t) : x \mapsto \gamma_x(t),$$

at time t, the velocity of the particle is $\gamma_x'(t) = \mathbf{v}(\gamma_x(t))$, the global velocity field is denoted as $\mathbf{v}(x, t)$, then the diffeomorphism satisfies the ODE:

$$\frac{d}{dt} g(x, t) = \mathbf{v}(g(x, t), t). \tag{10}$$

Diffeomorphism

Given a smooth velocity field $\mathbf{v}(x,t)$, we can get the diffeomorphism group $g(x,t)$. Namely, if $\mathbf{v}(x,t)$ is smooth enough, no shock waves will appear,

$$\partial_t \log\det\left[\frac{\partial g(x,t)}{\partial x}\right] = \nabla\cdot\mathbf{v}(g(x,t),t). \tag{11}$$

McCann Displacement

Definition (McCann's Displacement)

If the cost function is strictly convex, under Lagrangian point of view, all particles move with uniform speed in a straight line, their trajectories are

$$g_t(x) = (1-t)x + t(\nabla c)^{-1}(\nabla\varphi), \tag{12}$$

where φ is the optimal Kantorovich potential, this is called McCann displacement,

$$\rho_t = [(1-t)Id + t(\nabla c)^{-1}(\nabla\varphi)]_{\#}\mu.$$

McCann's displacement gives geodesics in Wasserstein space. One can show that

$$\mathcal{W}_c((g_s)_{\#}\mu, (g_t)_{\#}\mu) = |s-t|\mathcal{W}_c(\mu,\nu), \quad \forall s,t \in [0,1].$$

Time Dependent Optimal Transport Problem

Given a differential cost function $c(\mathbf{v})$, defined on velocity vector, then the cost for a trajectory is

$$\mathcal{C}[g_t(x)] := \int_0^1 c(\dot{g}_t(x))dt.$$

Problem (Time Dependent Optimal Transport)

Find a flow field connecting μ and ν, that minimizes the total cost of all trajectories:

$$\inf\left\{\int_\Omega \mathcal{C}[g_t(x)]d\mu(x) : \; g_0 = Id, \quad (g_1)_\# \mu = \nu\right\}. \tag{13}$$

Time Dependent Optimal Transport Problem

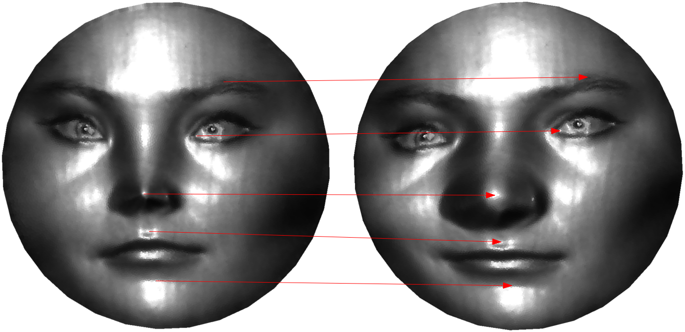

Figure: McCann's displacementEvery particle is in a uniform linear motion.

Time Dependent Optimal Transport Problem

From mass-conservation law, we get the continuity equation:

$$\frac{d}{dt}\rho(x,t) + \nabla \cdot (\rho(x,t)\mathbf{v}(x,t)) = 0,$$

By McCann interpolation, each trajectory is a constant velocity, the material derivative Dv/Dt is 0, hence the velocity field satisfies the Euler equation:

$$\frac{d}{dt}\mathbf{v}(x,t) + \mathbf{v}(x,t) \cdot \nabla \mathbf{v}(x,t) = 0.$$

The cost function information is implied by the initial conditon

$$\mathbf{v}(x,0) = (\nabla c)^{-1}(\nabla \varphi),$$

where φ is the optimal Kantorovich potential.

Benamou-Brenier Problem

Consider all flow fields connecting μ and ν, denote $(\rho(x,t), \mathbf{v}(x,t))$ as (ρ_t, v_t), then

$$\Gamma(\mu,\nu) := \left\{ (\rho_t, v_t) : \frac{\partial \rho_t}{\partial t} = -\nabla \cdot (\rho_t \mathbf{v}_t), \quad \rho_0 = \mu, \quad \rho_1 = \nu \right\}$$

Given any time $t \in [0,1]$, the Kinetic energy of the velocity field $\mathbf{v}(x,t)$ is defined as

$$E(\mathbf{v}_t) := \frac{1}{2} \int_\Omega \rho(x,t) \|\mathbf{v}(x,t)\|^2 dx.$$

Benamou-Brenier Problem

Problem (Benamou-Brenier)

Find the flow field in $\Gamma(\mu, \nu)$*, that minimizes the total kinetic energy,*

$$BB: \quad \min\left\{\frac{1}{2}\int_0^1 \int_\Omega \rho_t \|\mathbf{v}_t\|^2 dxdt: \ (\rho_t, \mathbf{v}_t) \in \Gamma(\mu, \nu)\right\} \tag{14}$$

Benamou-Brenier Problem

Use variational approach, assume $\rho_t \mathbf{w}_t$ is divergence free,

$$\frac{1}{2}\frac{d}{d\varepsilon}\int_0^1\int_\Omega \rho_t\langle \mathbf{v}_t+\varepsilon\mathbf{w}_t, \mathbf{v}_t+\varepsilon\mathbf{w}_t\rangle dxdt = \int_0^1\int_\Omega \rho_t\langle \mathbf{v}_t, \varepsilon\mathbf{w}_t\rangle dxdt = 0,$$

by Hodge decomposition theorem, $\mathbf{v}_t$ is orthogonal to all divergence free vector fields, so $\mathbf{v}_t = \nabla u_t$, where $u_t : \Omega \to \mathbb{R}$ is a family of funcitions. This can be obatined by McCann interpolation.

Benamou-Brenier Problem

Problem (Benamou-Brenier)

$$\mathcal{W}_2(\mu,\nu) := \min\left\{\int_0^1 \left\|\frac{\partial\rho}{\partial t}\right\|^2_{\rho(t)} dt, \rho_0=\mu, \rho_1=\nu, -\nabla\cdot(\rho\nabla u)=\frac{\partial\rho}{\partial t}\right\}$$

where

$$\left\|\frac{\partial\rho}{\partial t}\right\|^2_{\rho(t)} = \int_\Omega \rho|\nabla u|^2,$$

Otto's Interpretation

Given two geodesics $\rho_1(t), \rho_2(t) \subset \mathcal{P}(\Omega)$, $\rho_1(0) = \rho_2(0) = \rho$, the tangent vector at $\rho \in \mathcal{P}(\Omega)$,

$$\frac{\partial \rho_1}{\partial t} = -\nabla \cdot (\rho \nabla \varphi_1)$$
$$\frac{\partial \rho_2}{\partial t} = -\nabla \cdot (\rho \nabla \varphi_2)$$

the Riemannian inner product is

$$\left\langle \frac{\partial \rho_1}{\partial t}, \frac{\partial \rho_2}{\partial t} \right\rangle_\rho = \int_\Omega \rho \langle \nabla \varphi_1, \nabla \varphi_2 \rangle dx.$$

Entropy Flow

Definition (Entropy)

Given a probability measure $\rho \in \mathcal{P}(\Omega)$, its entropy is defined as

$$\mathrm{Ent}(\rho) := \int_\Omega \rho \log \rho dx.$$

Principle of maximum entropy

The principle of maximum entropy states that the probability distribution which best represents the current state of knowledge is the one with largest entropy, in the context of precisely stated prior data.

Entropy Flow

Consider a path $\rho(t) \subset \mathcal{P}(\Omega)$,

$$\frac{d}{dt}\text{Ent}(\rho(t)) = \int_\omega \left(\dot{\rho}\log\rho + \rho\frac{\dot{\rho}}{\rho}\right) dx = \int_\Omega (1+\log\rho)\dot{\rho}dx.$$

By continuity equation $\dot{\rho} = -\nabla\cdot(\rho\mathbf{v})$, assume $\Omega = \mathbb{R}^d$, hence

$$\int_\Omega \dot{\rho}dx = -\int_\Omega \nabla\cdot(\mathbf{v}\rho)dx = -\int_{\partial\Omega}\rho\mathbf{v}dx = 0.$$

We obtain

$$\frac{d}{dt}\text{Ent}(\rho(t)) = \int_\Omega \log\rho\dot{\rho}dx = -\int_\Omega \log\rho\nabla\cdot(\rho\mathbf{v}),$$

Entropy Flow

At the same time

$$\nabla \cdot (\rho \log \rho \mathbf{v}) = \log \rho \nabla \cdot (\rho \mathbf{v}) + \langle \nabla \log \rho, \rho \mathbf{v} \rangle,$$

We obtain

$$\begin{aligned} \frac{d}{dt}\text{Ent}(\rho(t)) &= -\int_\Omega \log \rho \nabla \cdot (\rho \mathbf{v}) \\ &= \int_\Omega \langle \nabla \log \rho, \rho \mathbf{v} \rangle - \int_{\partial\Omega} \rho \log \rho \mathbf{v} \\ &= \int_\Omega \langle \nabla \log \rho, \mathbf{v} \rangle \rho dx. \end{aligned}$$

This shows the Wasserstein gradient of Entropy is $\nabla \log \rho$.

Entropy Flow

In order to reduce the entropy, we let $\mathbf{v} = -\nabla \log \rho$, plug into the continuity equation:

$$\frac{\partial \rho_t}{\partial t} + \nabla \cdot (\rho_t \mathbf{v}_t) = 0,$$

hence

$$\begin{aligned} \frac{\partial \rho_t}{\partial t} - \nabla \cdot (\rho_t \nabla \log \rho_t) &= 0 \\ \frac{\partial \rho_t}{\partial t} - \nabla \cdot (\rho_t \frac{\nabla \rho_t}{\rho_t}) &= 0 \\ \frac{\partial \rho_t}{\partial t} - \Delta_t &= 0 \end{aligned}$$

This shows Wasserstein gradient flow of entropy equals to the classical heat flow.

Entropy Flow

We let $\mathbf{v} = -\nabla \log \rho$, plug into the continuity equation:

$$\frac{d}{dt}\mathrm{Ent}(\rho(t)) = \int_\Omega \langle \nabla \log \rho, \mathbf{v} \rangle \rho dx = -\int_\Omega \frac{|\nabla \rho|^2}{\rho} dx = -4\int_\Omega |\nabla \sqrt{\rho}|^2 dx.$$

This gives the dissipation speed of the entropy. Let t go to infinity, ρ_∞ becomes a uniform distribution.