Well-posedness and convergence of entropic approximation of semi-geostrophic equations

Hugo Malamut

Cérémade (Université Dauphine PSL)
Mokaplan (Inria)

March 25, 2025

Entropic Optimal Transport Solutions of the Semi-geostrophic equations

Colin Cotter — Imperial College
Jean-David Benamou, Hugo Malamut — MOKAPLAN, INRIA, PSL Dauphine, CNRS

Well-posedness and convergence of entropic approximation of semi-geostrophic equations

Guillaume Carlier, Hugo Malamut
MOKAPLAN
INRIA, PSL Dauphine, CNRS

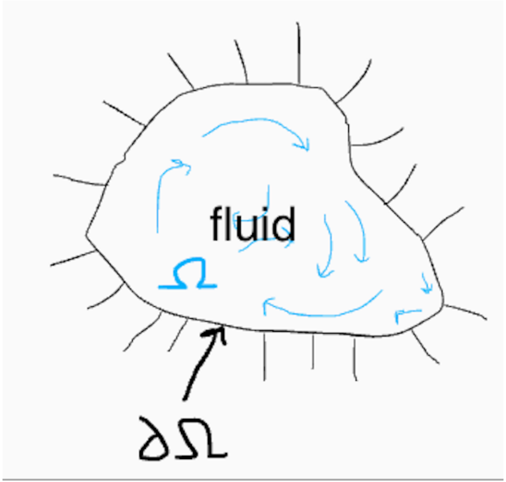

Figure: Fluid in a domain $\Omega \subset \mathbb{R}^n$, v_0 given

Euler equation

$$\underbrace{m\ddot{x}}_{\text{acceleration}} = \underbrace{\sum_i F_i}_{\text{forces}}$$

$$\underbrace{m\ddot{x}}_{\text{acceleration}} = \underbrace{\sum_i F_i}_{\text{forces}}$$

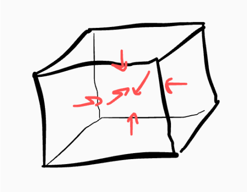

Figure: Pressure forces on a fluid parcel

$$F_p = pS\vec{n}_{int}$$

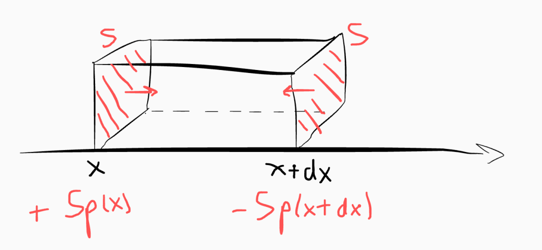

Figure: Horizontal pressure force: $F_{p,x} = -\partial_x p$

$$F_p = pS\vec{n}_{int}$$

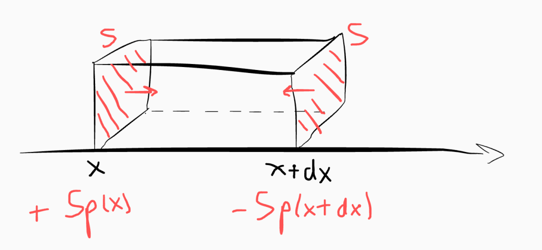

Figure: Horizontal pressure force: $F_{p,x} = -\partial_x p$

For a small enough parcel

$$F_p = -\nabla p$$

$$\underbrace{\ddot{x}}_{\text{acceleration}} = \underbrace{-\nabla p(x)}_{\text{pressure force}} + \underbrace{\sum_i F_i}_{\text{other forces}}$$

$$\underbrace{\ddot{x}}_{\text{acceleration}} = \underbrace{-\nabla p(x)}_{\text{pressure force}} + \underbrace{\sum_i F_i}_{\text{other forces}}$$

For Euler equation, there is no other force:

$$\ddot{x} = -\nabla p(x)$$

Let $X_t(a)$ or $X(t,a)$ be the resolvant of $\ddot{x} = -\nabla p_t(x)$:

$$\begin{cases} \frac{\partial^2}{\partial t^2} X(t,a) = -\nabla p_t(X(t,a)) \\ X(0,a) = a, \frac{\partial}{\partial t} X(t,a) = v_0(a) \end{cases}$$

Let $X_t(a)$ or $X(t,a)$ be the resolvant of $\ddot{x} = -\nabla p_t(x)$:

$$\begin{cases} \ddot{X}_t = -\nabla p_t \circ X_t \\ X_0 = id, \quad \dot{X}_0 = v_0 \end{cases}$$

Let $X_t(a)$ or $X(t,a)$ be the resolvant of $\ddot{x} = -\nabla p_t(x)$:

$$\begin{cases} \ddot{X}_t = -\nabla p_t \circ X_t \\ X_0 = id, \quad \dot{X}_0 = v_0 \end{cases}$$

Incompressibility constraint:

$$\forall t > 0 \quad (X_t)_\# \mathcal{L}_\Omega = \mathcal{L}_\Omega$$

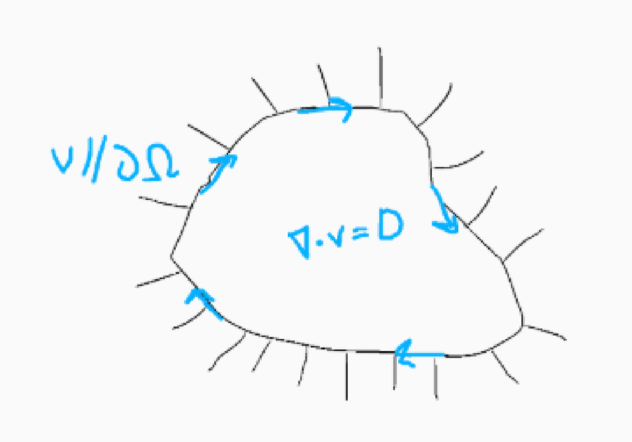

Figure: Conditions on the velocity field: $\nabla \cdot v = 0$ in Ω and $v//\partial\Omega$ on $\partial\Omega$

In Lagrangian coordinates:

$$\begin{cases} \ddot{X}_t = -\nabla p_t \circ X_t \\ (X_t)_{\#}\mathcal{L}_\Omega = \mathcal{L}_\Omega \end{cases}$$

with initial conditions $X_0 = id$, $\dot{X}_0 = v_0$

In Lagrangian coordinates:

$$\begin{cases} \ddot{X}_t = -\nabla p_t \circ X_t \\ (X_t)_\# \mathcal{L}_\Omega = \mathcal{L}_\Omega \end{cases}$$

with initial conditions $X_0 = id$, $\dot{X}_0 = v_0$
In Eulerian coordinates:

$$\begin{cases} \partial_t v + (v \cdot \nabla) v = -\nabla p \\ \nabla \cdot v = 0 \end{cases}$$

with initial condition $v(0, .) = v_0$

Arnold's interpretation

$$H = L^2(\Omega, \mathbb{R}^n)$$

Arnold's group is :

$$G = \{X : \Omega \to \Omega \,/\, X \text{ diffeomorphism and } X_{\#}\mathcal{L}_{\Omega} = \mathcal{L}_{\Omega}\}$$

$$H = L^2(\Omega, \mathbb{R}^n)$$

Arnold's group is :

$$G = \{X : \Omega \to \Omega \,/\, X \text{ diffeomorphism and } X_{\#}\mathcal{L}_\Omega = \mathcal{L}_\Omega\}$$

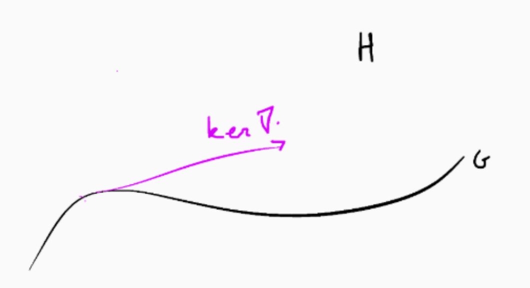

Figure: If $(X_t)_{\#}\mathcal{L} = \mathcal{L}$, $\dot{X}$ has 0 divergence: the tangent space is $ker(\nabla\cdot)$

Precise statement of Helmholtz decomposition:

$$\begin{array}{ccccc} C^1(\Omega, \mathbb{R}^n) & = & ker(\nabla \cdot) & \oplus^\perp & Im\nabla \\ v & = & w & + & \nabla p \end{array}$$

Precise statement of Helmholtz decomposition:

$$\begin{array}{ccccc} C^1(\Omega,\mathbb{R}^n) & = & ker(\nabla\cdot) & \oplus^\perp & Im\nabla \\ v & = & w & + & \nabla p \end{array}$$

Informal representation:

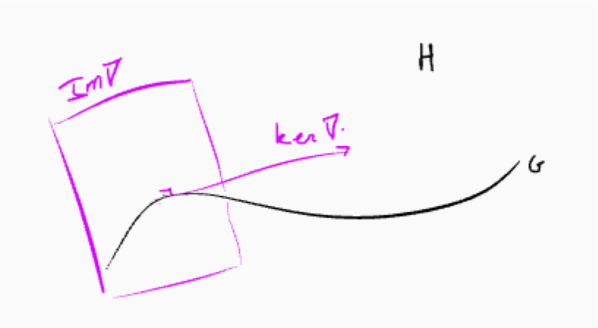

Figure: Formally, $T_X H = T_X G \oplus^\perp \{\nabla p, p \in C^2\}$

$$\ddot{X} = -\nabla p \circ X \Leftrightarrow \begin{array}{ccccc} T_X H & = & T_X G & \oplus^{\perp} & T_X G \\ \ddot{X} & = & 0 & + & -\nabla p \end{array}$$

$$\ddot{X} = -\nabla p \circ X \Leftrightarrow \begin{array}{ccccc} T_X H & = & T_X G & \oplus^\perp & T_X G \\ \ddot{X} & = & 0 & + & -\nabla p \end{array}$$

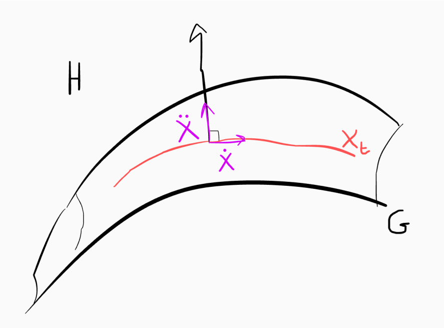

Figure: The acceleration is orthogonal to the manifold

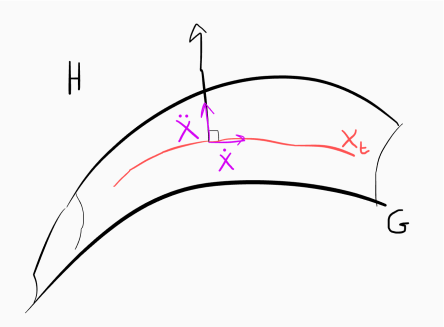

Figure: The acceleration is orthogonal to the manifold

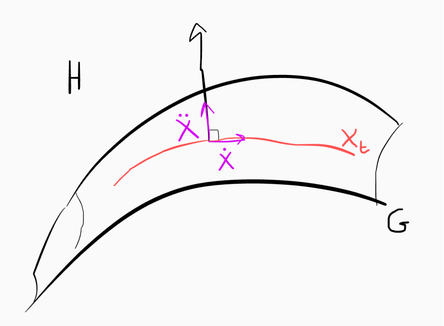

Figure: The acceleration is orthogonal to the manifold

$$\ddot{X} \perp T_X G$$

<u>X_t is a geodesic in G</u>

$$\ddot{X} \perp T_X G$$

X_t is a geodesic in G
Other interpretation: consider the problem

$$\min_{\substack{X_0, X_1 \, fixed \\ \dot{X} = v \circ X \\ \nabla \cdot v = 0}} \int_0^T \int_\Omega \frac{1}{2} |v|^2 \mathrm{d}a \mathrm{d}t$$

$$\ddot{X} \perp T_X G$$

X_t is a geodesic in G

Other interpretation: consider the problem

$$\min_{\substack{X_0, X_1 \, fixed \\ \dot{X}=v\circ X \\ \nabla\cdot v=0}} \int_0^T \int_\Omega \frac{1}{2}|v|^2 \mathrm{d}a\mathrm{d}t$$

Formally, the solution follows the Euler equation and the pressure is the Lagrange multiplier for the constraint $(X_t)_\# \mathcal{L}_\Omega = \mathcal{L}_\Omega$

Numerics: Optimal transport method for Euler equation (Cauchy problem)

Figure: The initial value (Cauchy) problem

Figure: The initial value (Cauchy) problem

The problem:

- we want to construct $X_t \in G$
- we can construct orthogonal acceleration only if for $Y \notin G$
- if $Y \notin G$, there is no equation on Y (!)

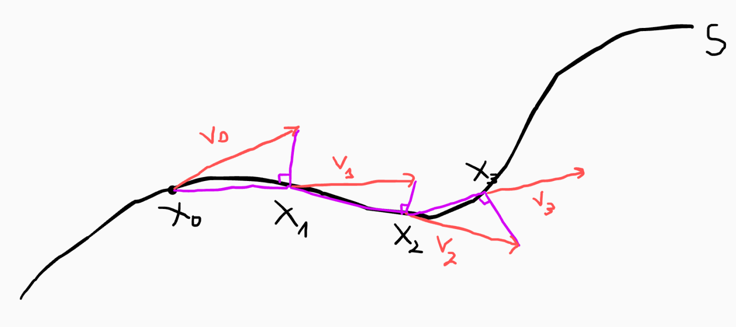

Figure: Project at each step

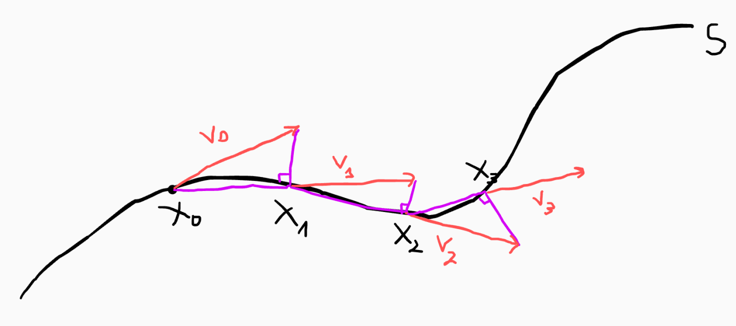

Figure: Project at each step

- original idea from Brenier
- no proof of convergence
- only discret in time

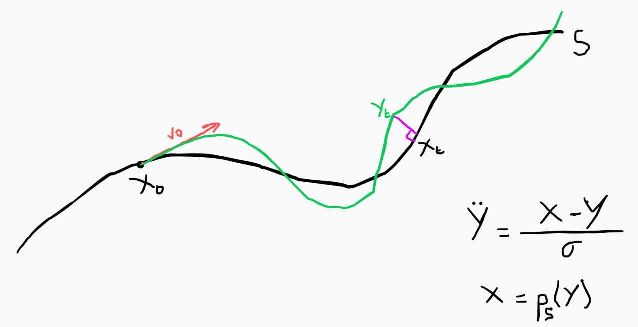

Figure: Allow oscillations with parameter σ

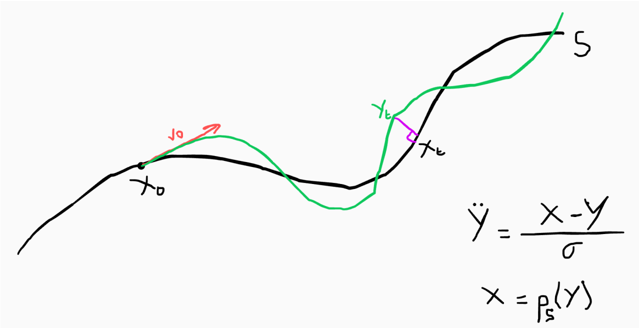

Figure: Allow oscillations with parameter σ

- original idea from Brenier-Loeper
- continuous time
- proof of convergence

How to project on G?

How to project on G?

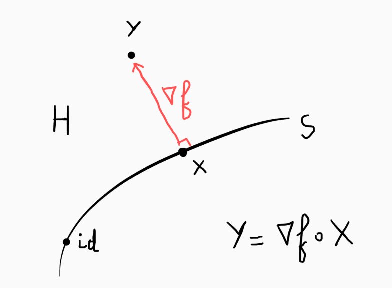

Figure: The polar factorisation theorem

We need to solve for the optimal transport problem between $\mathcal{L}_\Omega$ and $\mu := Y_\# \mathcal{L}_\Omega$:

$$\inf_{T_\# \mu = \mathcal{L}_\Omega} \int |T(a) - a|^2 d\mu(a) =? \inf_{T_\# \mathcal{L}_\Omega = \mu} \int_\Omega |T(a) - a|^2 da$$

We need to solve for the optimal transport problem between $\mathcal{L}_\Omega$ and $\mu := Y_\# \mathcal{L}_\Omega$:

$$\inf_{T_\# \mu = \mathcal{L}_\Omega} \int |T(a) - a|^2 d\mu(a) =? \inf_{T_\# \mathcal{L}_\Omega = \mu} \int_\Omega |T(a) - a|^2 da$$

- The optimizers are $T = \nabla f$ such that $\mu = T_\# \mathcal{L}_\Omega$ and $T^{-1} = \nabla f^*$ s.t. $\mathcal{L}_\Omega = (T^{-1})_\# \mu$
- $X = Y \circ \nabla f^*$ is the projection of Y on G

Semigeostrophic equations: Coriolis force's entrance

$$\underbrace{\ddot{x}}_{\text{acceleration}} = \underbrace{-\nabla p(x)}_{\text{pressure force}} + \underbrace{\sum_i F_i}_{\text{other forces}}$$

$$\underbrace{\ddot{x}}_{\text{acceleration}} = \underbrace{-\nabla p(x)}_{\text{pressure force}} + \underbrace{\sum_i F_i}_{\text{other forces}}$$

Geostrophic regime: only one other force, the Coriolis force

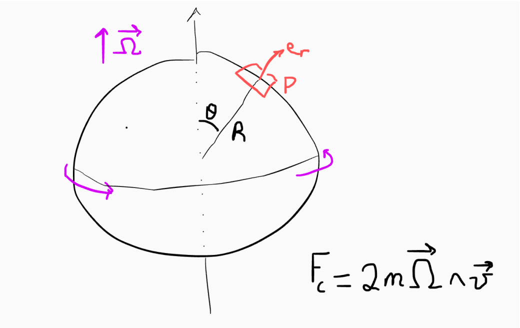

Figure: the rotation of the earth, we decompose $\vec{\Omega} = \Omega_P + \Omega_r$

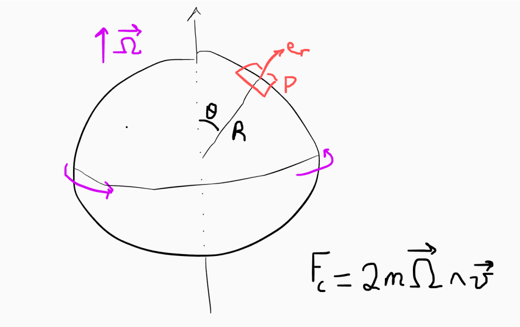

Figure: the rotation of the earth, we decompose $\vec{\Omega} = \Omega_P + \Omega_r$

- the vertical component $(F_c \cdot e_r)e_r$ is negligible with respect to gravity
- the speed is horizontal $(v \in P)$ so $\Omega_P \wedge v // e_r$

$$F_c = 2m(\cancel{\Omega_P} + \Omega_r) \wedge v$$

$$F_c = 2m(\not{\Omega_P} + \Omega_r) \wedge v$$

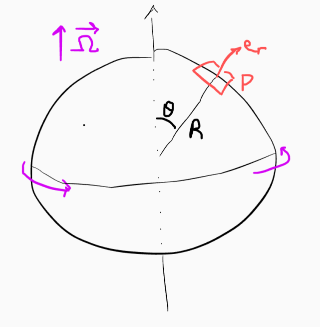

Figure: The vertical part of $\vec{\Omega}$ is $\Omega_r = \Omega R \cos(\theta) e_r$

$$F_c = 2m(\Omega_P + \Omega_r) \wedge v$$

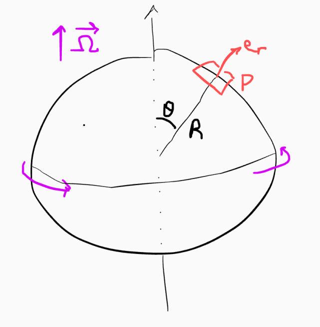

Figure: The vertical part of $\vec{\Omega}$ is $\Omega_r = \Omega R \cos(\theta) e_r$

Let $f = 2\Omega R \cos(\theta)$ The Coriolis force is $F_c = f e_r \wedge v$

The Coriolis force is $F_c = fe_r \wedge v$.

The Coriolis force is $F_c = fe_r \wedge v$.

Let $J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. In the horizontal coordinates, $e_r \wedge v = -Jv$

So the Coriolis force is $F_c = mfJv$

The Coriolis force is $F_c = fe_r \wedge v$.
Let $J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. In the horizontal coordinates, $e_r \wedge v = -Jv$
So the Coriolis force is $F_c = mfJv$
Hence

$$\underbrace{\ddot{x}}_{\text{acceleration}} = \underbrace{-\nabla p(x)}_{\text{pressure}} \underbrace{-fJ\dot{x}}_{\text{Coriolis}}$$

The Coriolis force is $F_c = fe_r \wedge v$.
Let $J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. In the horizontal coordinates, $e_r \wedge v = -Jv$
So the Coriolis force is $F_c = mfJv$
Hence

$$\underbrace{\ddot{x}}_{\text{acceleration}} = \underbrace{-\nabla p(x)}_{\text{pressure}} \underbrace{-fJ\dot{x}}_{\text{Coriolis}}$$

Or

$$\ddot{x} + fJ\dot{x} + \nabla p(x) = 0$$

From now on, we suppose that f does not vary two much and we set $f = 1$

Cullen's stability principle

Take two parcels x and y with close initial conditions

$$\ddot{y} + J\dot{y} + \nabla p(y) = 0$$
$$\ddot{x} + J\dot{x} + \nabla p(x) = 0$$

Take two parcels x and y with close initial conditions

$$\ddot{y} + J\dot{y} + \nabla p(y) = 0$$
$$\ddot{x} + J\dot{x} + \nabla p(x) = 0$$

Set $r = y - x$ and approximate $\nabla p(y) - \nabla p(x) \simeq D^2 p(x)(y - x)$

$$\ddot{r} + J\dot{r} + D^2 p(x) r = 0$$

Consider time scale under which $D^2 p(x)$ does not vary too much,

$$\ddot{r} + J\dot{r} + D^2 p(x) r = 0$$

$$\ddot{r} + J\dot{r} + D^2 p(x) r = 0$$

- $D^2 p(x)$ is a real symmetric

$$\ddot{r} + J\dot{r} + D^2 p(x) r = 0$$

- $D^2 p(x)$ is a real symmetric
- J is orthogonal

$$\ddot{r} + J\dot{r} + D^2 p(x) r = 0$$

- $D^2 p(x)$ is a real symmetric
- J is orthogonal
- we can suppose $D^2 p(x)$ diagonal: $D^2 p(x) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$

$$\ddot{r} + J\dot{r} + D^2p(x)r = 0$$

- $D^2p(x)$ is a real symmetric
- J is orthogonal
- we can suppose $D^2p(x)$ diagonal: $D^2p(x) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$

$$\ddot{r} + J\dot{r} + D^2p(x)r = 0$$

- $D^2p(x)$ is a real symmetric
- J is orthogonal
- we can suppose $D^2p(x)$ diagonal: $D^2p(x) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$

We can rewrite

$$\begin{pmatrix} \ddot{r} \\ \dot{r} \end{pmatrix} = \begin{pmatrix} J & -D^2p(x) \\ I & 0 \end{pmatrix} \begin{pmatrix} \dot{r} \\ r \end{pmatrix}.$$

$$\ddot{r} + J\dot{r} + D^2p(x)r = 0$$

- $D^2p(x)$ is a real symmetric
- J is orthogonal
- we can suppose $D^2p(x)$ diagonal: $D^2p(x) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$

We can rewrite

$$\begin{pmatrix} \ddot{r} \\ \dot{r} \end{pmatrix} = \begin{pmatrix} J & -D^2p(x) \\ I & 0 \end{pmatrix} \begin{pmatrix} \dot{r} \\ r \end{pmatrix}.$$

What matters for stability are the eigenvalues of

$$A = \begin{pmatrix} J & -D^2p(x) \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -a & 0 \\ 1 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

What matters for stability are the eigenvalues of

$$A = \begin{pmatrix} J & -D^2 p(x) \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -a & 0 \\ 1 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Facts:

- $$\chi_A(\lambda) = \lambda^4 + (a + b - 1)\lambda^2 - ab$$

What matters for stability are the eigenvalues of

$$A = \begin{pmatrix} J & -D^2 p(x) \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -a & 0 \\ 1 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Facts:

- $$\chi_A(\lambda) = \lambda^4 + (a + b - 1)\lambda^2 - ab$$
- λ eigenvalue $\Leftrightarrow$ $-\lambda$ eigenvalue

What matters for stability are the eigenvalues of

$$A = \begin{pmatrix} J & -D^2p(x) \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -a & 0 \\ 1 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Facts:

- $$\chi_A(\lambda) = \lambda^4 + (a+b-1)\lambda^2 - ab$$
- λ eigenvalue $\Leftrightarrow$ $-\lambda$ eigenvalue
- $Re(\lambda) \geq 0$ and $Re(-\lambda) \geq 0$ implies $Re(\lambda) = 0$ and $\lambda^2 \in \mathbb{R}^{-1}$

What matters for stability are the eigenvalues of

$$A = \begin{pmatrix} J & -D^2p(x) \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -a & 0 \\ 1 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Facts:

- $$\chi_A(\lambda) = \lambda^4 + (a+b-1)\lambda^2 - ab$$
- λ eigenvalue $\Leftrightarrow$ $-\lambda$ eigenvalue
- $Re(\lambda) \geq 0$ and $Re(-\lambda) \geq 0$ implies $Re(\lambda) = 0$ and $\lambda^2 \in \mathbb{R}^{-1}$
- $\mu = \lambda^2$ solves
$$\mu^2 - (a+b+1)\mu + ab = 0$$

What matters for stability are the eigenvalues of

$$A = \begin{pmatrix} J & -D^2 p(x) \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -a & 0 \\ 1 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Facts:

- $$\chi_A(\lambda) = \lambda^4 + (a+b-1)\lambda^2 - ab$$
- λ eigenvalue $\Leftrightarrow$ $-\lambda$ eigenvalue
- $Re(\lambda) \geq 0$ and $Re(-\lambda) \geq 0$ implies $Re(\lambda) = 0$ and $\lambda^2 \in \mathbb{R}^{-1}$
- $\mu = \lambda^2$ solves
$$\mu^2 - (a+b+1)\mu + ab = 0$$
- we need
$$a+b-1 \geq 0 \quad \text{and} \quad ab \geq 0$$

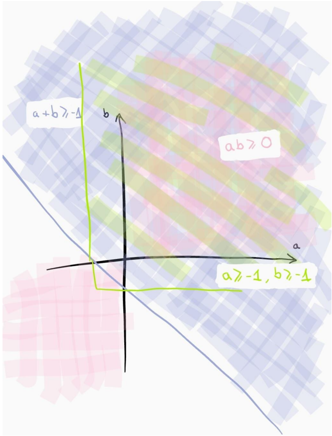

Figure: We have $a \geq -1$ and $b \geq -1$

The eigenvalues of $D^2 p(x)$ are larger than -1,

The eigenvalues of $D^2 p(x)$ are larger than -1,
Equivalently $I + D^2 p(x)$ is semidefinite positive,

The eigenvalues of $D^2 p(x)$ are larger than -1,
Equivalently $I + D^2 p(x)$ is semidefinite positive,
Equivalently

$$x \mapsto x + \nabla p(x) \text{ is the gradient of a convex function}$$

The eigenvalues of $D^2 p(x)$ are larger than -1,
Equivalently $I + D^2 p(x)$ is semidefinite positive,
Equivalently

$$x \mapsto x + \nabla p(x) \text{ is the gradient of a convex function}$$

To be continued...

Semigeostrophic approximation

$T := t_1 - t_0 >> 1$ the time scale, we are interested in

$$y(t) := x(t_0 + tT).$$

We still have

$$\ddot{x} + J\dot{x} + \nabla p(x) = 0$$

$T := t_1 - t_0 >> 1$ the time scale, we are interested in

$$y(t) := x(t_0 + tT).$$

We still have

$$\ddot{x} + J\dot{x} + \nabla p(x) = 0$$

In terms of y, we have

$$\frac{1}{T^2}\ddot{y} + \frac{1}{T}J\dot{y} + \nabla p(y) = 0$$

First step:

$$\cancel{\frac{1}{T^2}\ddot{y}} + \frac{1}{T} J\dot{y} + \nabla p(y) = 0,$$

So (remember that $J^{-1} = -J$):

First step:

$$\cancel{\frac{1}{T^2}\ddot{y}} + \frac{1}{T} J\dot{y} + \nabla p(y) = 0,$$

So (remember that $J^{-1} = -J$):

$$\dot{y} = \underbrace{TJ\nabla p(y)}_{\text{Geostrophic wind}}$$

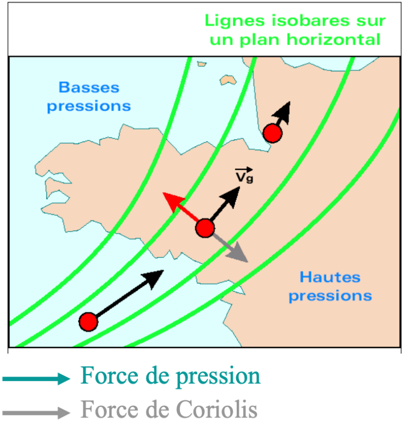

Figure: Excerpt from [?], wind map here

Semi-geostrophic approximation:

$$\frac{1}{T^2}\ddot{y} + \frac{1}{T}J\dot{y} + \nabla p(y) = 0$$

Second step: we suppose

$$\dot{y} = \underbrace{TJ\nabla p(y)}_{\text{Geostrophic wind}} + \underbrace{v_a}_{\text{ageostrophic speed}} .$$

So

$$\frac{1}{T^2}\dot{v_a} + \frac{1}{T}J\nabla \dot{p}(y) + \frac{1}{T}J\dot{y} + \nabla p(y) = 0$$

Semi-geostrophic approximation:

$$\frac{1}{T^2}\ddot{y} + \frac{1}{T}J\dot{y} + \nabla p(y) = 0$$

Second step: we suppose

$$\dot{y} = \underbrace{TJ\nabla p(y)}_{\text{Geostrophic wind}} + \underbrace{v_a}_{\text{ageostrophic speed}} .$$

So

$$\cancel{\frac{1}{T^2}}\dot{v}_a + \frac{1}{T}J\nabla \dot{p}(y) + \frac{1}{T}J\dot{y} + \nabla p(y) = 0$$

Semi-geostrophic approximation:

$$\frac{1}{T^2}\ddot{y} + \frac{1}{T} J\dot{y} + \nabla p(y) = 0$$

Second step: we suppose

$$\dot{y} = \underbrace{TJ\nabla p(y)}_{\text{Geostrophic wind}} + \underbrace{v_a}_{\text{ageostrophic speed}} .$$

So

$$\cancel{\frac{1}{T^2}}\dot{v}_a + \frac{1}{T} J\nabla \dot{p}(y) + \frac{1}{T} J\dot{y} + \nabla p(y) = 0$$

We arrive to

$$\dot{y} + \underbrace{\nabla \dot{p}(y)}_{\text{ageostrophic correction}} = \underbrace{TJ\nabla p(y)}_{\text{geostrophic wind}}$$

x_t and p_t solves the semigeostrophic equation if

$$\dot{x} + \dot{\nabla p}(x) = J\nabla p(x)$$

- $x_g := x + \nabla p(x)$ is called geostrophic coordinates
- $v_g := J\nabla p$ is called the geostrophic wind

We have

$$\dot{X}_g = v_g \circ X$$

We have

$$\dot{X}_g = v_g \circ X$$

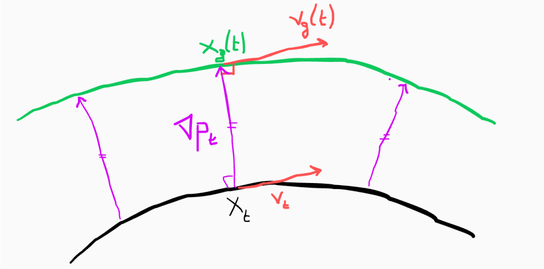

Figure: The speed of X_g is $v_g = J\nabla p$

We have

$$\dot{X}_g = v_g \circ X$$

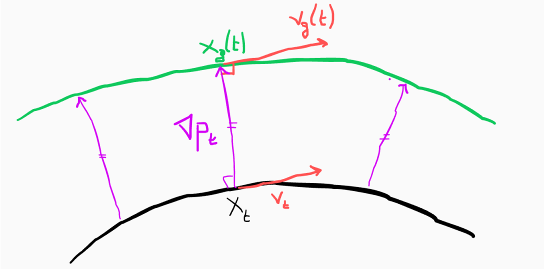

Figure: The speed of X_g is $v_g = J\nabla p$

- we have an equation on X_g!
- The energy $E = \int_\Omega \frac{1}{2}|\dot{X}_g|^2 da = \frac{1}{2}\|\nabla p\|^2_{L^2(\Omega)}$ is conserved

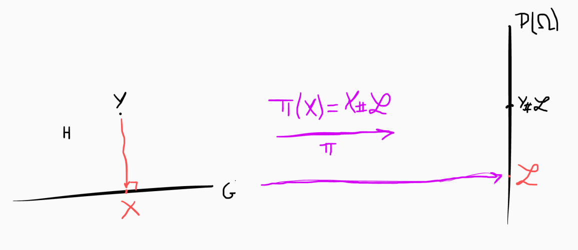

Figure: Since we have a equation on $X_g \notin G$, we can project the equation on the measures

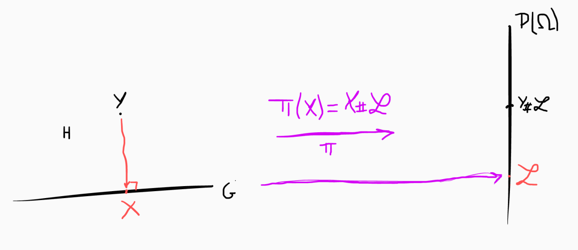

Figure: Since we have a equation on $X_g \notin G$, we can project the equation on the measures

Definition (Semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t = -\nabla \cdot [\rho_t J(T^*_{\rho_t} - id)] \\ T^*_{\rho_t} \text{ optimal map between } \rho_t \text{ and } \mathcal{L}_\Omega \end{cases} \tag{SG}$$

- Euler equation:
  [Arnold, 1966],[Brenier and Loeper, 2004],[Gallouët and Mérigot, 2018],[De Goes et al., 2015]
- Hoskins' transform [Hoskins, 1975]
- Cullen's stability principle [Cullen and Purser, 1984]
- Link with the optimal transport, weak solution in geostrophic coordinates
  [Benamou and Brenier, 1998]
- A litterature on interpreting these in physical coordinates [Figalli, 2018].
- Hamiltonian structure [Ambrosio and Gangbo, 2008]
- Numerics : Semi-Discrete optimal transport [Bourne et al., 2022]

Entropic optimal transport resolution

Relative entropy :

$$\mathcal{H}(\mu|\nu) := \begin{cases} \int \log(\frac{d\mu}{d\nu})\, d\mu \text{ if } \mu << \nu \\ +\infty \text{ else} \end{cases}$$

Relative entropy :

$$\mathcal{H}(\mu|\nu) := \begin{cases} \int \log(\frac{d\mu}{d\nu})\, d\mu \text{ if } \mu << \nu \\ +\infty \text{ else} \end{cases}$$

Definition (Entropic optimal transport)

$$OT_\varepsilon(\mu,\nu) = \min_{\pi\in\Pi(\mu,\nu)} \int |x-g|^2 d\pi(x,g) + \varepsilon\mathcal{H}(\pi|\mu\otimes\nu)$$

If π_ε is optimal, the <u>ε-Brenier map</u> is $T^\varepsilon(x) := \int g\pi_\varepsilon^x(dg)$[1]

[1] π_ε^x is the conditional distribution of π_ε

Relative entropy :

$$\mathcal{H}(\mu|\nu) := \begin{cases} \int \log(\frac{d\mu}{d\nu})\, d\mu \text{ if } \mu << \nu \\ +\infty \text{ else} \end{cases}$$

Definition (Entropic optimal transport)

$$OT_\varepsilon(\mu,\nu) = \min_{\pi\in\Pi(\mu,\nu)} \int |x-g|^2 d\pi(x,g) + \varepsilon\mathcal{H}(\pi|\mu\otimes\nu)$$

If π_ε is optimal, the ε-Brenier map is $T^\varepsilon(x) := \int g\pi_\varepsilon^x(dg)$[1]

Advantage : Sinkhorn algorithm

[1] π_ε^x is the conditional distribution of π_ε

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t = -\nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] \\ T^{\varepsilon}_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L}_\Omega \end{cases} \tag{SG_ε}$$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t = -\nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] \\ T^{\varepsilon}_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L}_\Omega \end{cases} \qquad (SG_\varepsilon)$$

- Space discretization : $\rho_t \simeq \sum_{i=1}^{N} a_i \delta_{x_i(t)}$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t = -\nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] \\ T^{\varepsilon}_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L}_\Omega \end{cases} \qquad (SG_\varepsilon)$$

- Space discretization : $\rho_t \simeq \sum_{i=1}^N a_i \delta_{x_i(t)}$
- Euler scheme : $x_i(t+dt) \simeq x_i(t) + dtJ\left[T^{\varepsilon}_{\rho_t}(x_i(t)) - x_i(t)\right]$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t = -\nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] \\ T^{\varepsilon}_{\rho_t}\ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L}_\Omega \end{cases} \qquad (SG_\varepsilon)$$

- Space discretization : $\rho_t \simeq \sum_{i=1}^{N} a_i \delta_{x_i(t)}$
- Euler scheme : $x_i(t+dt) \simeq x_i(t) + dtJ\left[T^{\varepsilon}_{\rho_t}(x_i(t)) - x_i(t)\right]$
- $T^{\varepsilon}_{\rho_t}$ using Geomloss (J. Feydy)
  `https://www.kernel-operations.io/geomloss/` A fast GPU based Sinkhorn solver.

Link to the movie : https://www.youtube.com/watch?v=U2j_AY-2FbM

Well-posedness of entropic semi-geostrophic equations ($\varepsilon > 0$)

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t + \nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] = 0 \\ T^{\varepsilon}_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L} \end{cases} \tag{SG_ε}$$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t + \nabla \cdot [\rho_t J(T^\varepsilon_{\rho_t} - id)] = 0 \\ T^\varepsilon_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L} \end{cases} \qquad (SG_\varepsilon)$$

- Flow X_t: given V and ρ_0, computes ρ_t s.t.

$$\partial_t \rho + \nabla \cdot (\rho J V) = 0$$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t + \nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] = 0 \\ T^{\varepsilon}_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L} \end{cases} \tag{SG_ε}$$

- Flow X_t: given V and ρ_0, computes ρ_t s.t.

$$\partial_t \rho + \nabla \cdot (\rho J V) = 0$$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t + \nabla \cdot [\rho_t J(T^\varepsilon_{\rho_t} - id)] = 0 \\ T^\varepsilon_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L} \end{cases} \quad (SG_\varepsilon)$$

- Flow X_t: given V and ρ_0, computes ρ_t s.t.

$$\partial_t \rho + \nabla \cdot (\rho J V) = 0$$

"$\rho_t = V \cdot \rho_0$"

- entropic OT: given ρ computes $V_\varepsilon = T^\varepsilon_\rho - id$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t + \nabla \cdot [\rho_t J(T^\varepsilon_{\rho_t} - id)] = 0 \\ T^\varepsilon_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L} \end{cases} \quad (SG_\varepsilon)$$

- Flow X_t: given V and ρ_0, computes ρ_t s.t.

$$\partial_t \rho + \nabla \cdot (\rho J V) = 0$$

$$\text{"}\rho_t = V \cdot \rho_0\text{"}$$

- entropic OT: given ρ computes $V_\varepsilon = T^\varepsilon_\rho - id$

Definition (Entropic semi-geostrophic equations)

$$\begin{cases} \partial_t \rho_t + \nabla \cdot [\rho_t J(T^{\varepsilon}_{\rho_t} - id)] = 0 \\ T^{\varepsilon}_{\rho_t} \ \varepsilon\text{-Brenier map between } \rho_t \text{ and } \mathcal{L} \end{cases} \tag{SG_ε}$$

- Flow X_t: given V and ρ_0, computes ρ_t s.t.

$$\partial_t \rho + \nabla \cdot (\rho J V) = 0$$

$$"\rho_t = V \cdot \rho_0"$$

- entropic OT: given ρ computes $V_\varepsilon = T^{\varepsilon}_{\rho} - id$

$$"V_\varepsilon = V_\varepsilon[\rho]"$$

Definition (Schrödinger map)

Fix $\mu_0 \in \mathcal{P}(\Omega)$. We define the Schrödinger map $V_\varepsilon : \mathcal{P}(\Omega) \to L^\infty(G)$ by

$$\forall \rho, g \in \mathcal{P}(\mathcal{G}) \times \mathcal{G} \quad V_\varepsilon[\rho](g) = T_\rho^\varepsilon(g) - g$$

where T_ρ^ε is the ε-Brenier map between ρ and μ_0 ($T^\varepsilon(g) := \int x \pi_\varepsilon^g(dx)$)

ρ_t is a solution of (SG_ε) if and only if[2] $\rho_t = V_\varepsilon[\rho_t] \cdot \rho_0$

Proposition

If $\rho_0 \in \mathcal{P}(\mathcal{B}_R)$ *then* $\rho_t \in \mathcal{P}(B_{R_T})$ *with* $R_T := 2R_0 e^T$. *Define* $K := K_{R_T}$ *and* $M = M_{R_T}$.

[2] recall that $\rho_t = v \cdot \rho_0$ is equivalent to $\partial\rho + \nabla \cdot (\rho J v) = 0$

ρ_t is a solution of (SG_ε) if and only if[2] $\rho_t = V_\varepsilon[\rho_t] \cdot \rho_0$

Proposition

If $\rho_0 \in \mathcal{P}(\mathcal{B}_R)$ then $\rho_t \in \mathcal{P}(B_{R_T})$ with $R_T := 2R_0 e^T$. Define $K := K_{R_T}$ and $M = M_{R_T}$.
$\rho_t \mapsto V_\varepsilon[\rho_t] \cdot \rho_0$ is a contraction for the distance

$$dist(\rho^1, \rho^2) := \sup_{t\in[0,T]} e^{-\lambda t} W_2(\rho_t^1, \rho_t^2)$$

with $\lambda = \frac{5M}{2} + K$.

$$dist(V_\varepsilon[\rho_t^1] \cdot \rho_0, V_\varepsilon[\rho_t^2] \cdot \rho_0) \leq \frac{1}{2} dist(\rho^1, \rho^2)$$

[2] recall that $\rho_t = v \cdot \rho_0$ is equivalent to $\partial\rho + \nabla \cdot (\rho J v) = 0$

ρ_t is a solution of (SG_ε) if and only if[2] $\rho_t = V_\varepsilon[\rho_t] \cdot \rho_0$

Proposition

If $\rho_0 \in \mathcal{P}(\mathcal{B}_R)$ then $\rho_t \in \mathcal{P}(B_{R_T})$ with $R_T := 2R_0 e^T$. Define $K := K_{R_T}$ and $M = M_{R_T}$.
$\rho_t \mapsto V_\varepsilon[\rho_t] \cdot \rho_0$ is a contraction for the distance

$$dist(\rho^1, \rho^2) := \sup_{t \in [0,T]} e^{-\lambda t} W_2(\rho^1_t, \rho^2_t)$$

with $\lambda = \frac{5M}{2} + K$.

$$dist(V_\varepsilon[\rho^1_t] \cdot \rho_0, V_\varepsilon[\rho^2_t] \cdot \rho_0) \leq \frac{1}{2} dist(\rho^1, \rho^2)$$

(SG_ε) has a unique weak solution

[2] recall that $\rho_t = v \cdot \rho_0$ is equivalent to $\partial \rho + \nabla \cdot (\rho J v) = 0$

Convergence of entropic semi-geostrophic equations

weak solution of (SG_ε) on $[0, T] \times \mathbb{R}^d$

- for every $f \in C_c^1([0, T] \times \mathbb{R}^d)$, one has

$$\int_0^T \int_{\mathbb{R}^d} [\partial_t f + Jx \cdot \nabla f] \rho_t(\mathrm{d}x)\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d \times \mathbb{R}^d} Jy \cdot \nabla f(t, x) \gamma_t^\varepsilon(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t$$
$$= \int_{\mathbb{R}^d} f(T, x)\rho_T(\mathrm{d}x) - \int_{\mathbb{R}^d} f(0, x)\rho_0(\mathrm{d}x), \quad (1)$$

weak solution of (SG_ε) on $[0, T] \times \mathbb{R}^d$

- for every $f \in C_c^1([0, T] \times \mathbb{R}^d)$, one has

$$\int_0^T \int_{\mathbb{R}^d} [\partial_t f + Jx \cdot \nabla f] \rho_t(\mathrm{d}x)\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d \times \mathbb{R}^d} Jy \cdot \nabla f(t,x) \gamma_t^\varepsilon(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t$$
$$= \int_{\mathbb{R}^d} f(T,x)\rho_T(\mathrm{d}x) - \int_{\mathbb{R}^d} f(0,x)\rho_0(\mathrm{d}x), \quad (1)$$

- γ_t^ε is t-a.e. an ε-optimal plan between ρ_t and $\mathcal{L}_\Omega$

weak solution of (SG_ε) on $[0, T] \times \mathbb{R}^d$

- for every $f \in C_c^1([0, T] \times \mathbb{R}^d)$, one has

$$\int_0^T \int_{\mathbb{R}^d} [\partial_t f + Jx \cdot \nabla f] \rho_t(\mathrm{d}x)\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d \times \mathbb{R}^d} Jy \cdot \nabla f(t,x) \gamma_t^\varepsilon(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t$$
$$= \int_{\mathbb{R}^d} f(T,x)\rho_T(\mathrm{d}x) - \int_{\mathbb{R}^d} f(0,x)\rho_0(\mathrm{d}x), \quad (1)$$

- γ_t^ε is t-a.e. an ε-optimal plan between ρ_t and $\mathcal{L}_\Omega$

weak solution of (SG_ε) on $[0, T] \times \mathbb{R}^d$

- for every $f \in C_c^1([0, T] \times \mathbb{R}^d)$, one has

$$\int_0^T \int_{\mathbb{R}^d} [\partial_t f + Jx \cdot \nabla f]\rho_t(\mathrm{d}x)\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d \times \mathbb{R}^d} Jy \cdot \nabla f(t,x)\gamma_t^\varepsilon(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t$$
$$= \int_{\mathbb{R}^d} f(T,x)\rho_T(\mathrm{d}x) - \int_{\mathbb{R}^d} f(0,x)\rho_0(\mathrm{d}x), \quad (1)$$

- γ_t^ε is t-a.e. an ε-optimal plan between ρ_t and $\mathcal{L}_\Omega$

Strategy :

- extract a subsequence $\rho_t^{\varepsilon_n}$

weak solution of (SG_ε) on $[0, T] \times \mathbb{R}^d$

- for every $f \in C_c^1([0, T] \times \mathbb{R}^d)$, one has

$$\int_0^T \int_{\mathbb{R}^d} [\partial_t f + Jx \cdot \nabla f] \rho_t(\mathrm{d}x)\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d \times \mathbb{R}^d} Jy \cdot \nabla f(t, x) \gamma_t^\varepsilon(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t = \int_{\mathbb{R}^d} f(T, x)\rho_T(\mathrm{d}x) - \int_{\mathbb{R}^d} f(0, x)\rho_0(\mathrm{d}x), \quad (1)$$

- γ_t^ε is t-a.e. an ε-optimal plan between ρ_t and $\mathcal{L}_\Omega$

Strategy :

- extract a subsequence $\rho_t^{\varepsilon_n}$
- Pass to the limit

- Semi-geostrophic shallow water :
  $\mathcal{L}_\Omega$ replaced by $h_t(x)d\mathcal{L}_\Omega$
  h_t : depth of water solves

$$\inf_h OT_0(\rho_t, h) + \int h^2 d\mathcal{L}_\Omega$$

- Compressible semi-geostrophic : the cost is

$$\frac{1}{Z}(|x - X|^2 - z)$$

- 3D...

Thank you!

Ambrosio, L. and Gangbo, W. (2008). Hamiltonian odes in the wasserstein space of probability measures. Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, 61(1):18–53.

Arnold, V. (1966). Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In Annales de l'institut Fourier, volume 16, pages 319–361.

Benamou, J.-D. and Brenier, Y. (1998). Weak existence for the semigeostrophic equations formulated as a coupled monge–ampere/transport problem. SIAM Journal on Applied Mathematics, 58(5):1450–1461.

Bourne, D. P., Egan, C. P., Pelloni, B., and Wilkinson, M. (2022). Semi-discrete optimal transport methods for the semi-geostrophic equations. Calculus of Variations and Partial Differential Equations, 61:1–34.

Brenier, Y. and Loeper, G. (2004). A geometric approximation to the euler equations: the vlasov–monge–ampere system. Geometric & Functional Analysis GAFA, 14:1182–1218.

Cullen, M. J. and Purser, R. J. (1984). An extended lagrangian theory of semi-geostrophic frontogenesis. Journal of Atmospheric Sciences, 41(9):1477–1497.

De Goes, F., Wallez, C., Huang, J., Pavlov, D., and Desbrun, M. (2015). Power particles: an incompressible fluid solver based on power diagrams. ACM Trans. Graph., 34(4):50–1.

Figalli, A. (2018). Global Existence for the Semigeostrophic Equations via Sobolev Estimates for Monge-Ampère, pages 1–42. Springer International Publishing, Cham.

Gallouët, T. O. and Mérigot, Q. (2018). A lagrangian scheme à la brenier for the incompressible euler equations. Foundations of Computational Mathematics, 18(4):835–865.

Hoskins, B. J. (1975). The geostrophic momentum approximation and the semi-geostrophic equations. Journal of Atmospheric Sciences, 32(2):233–242.

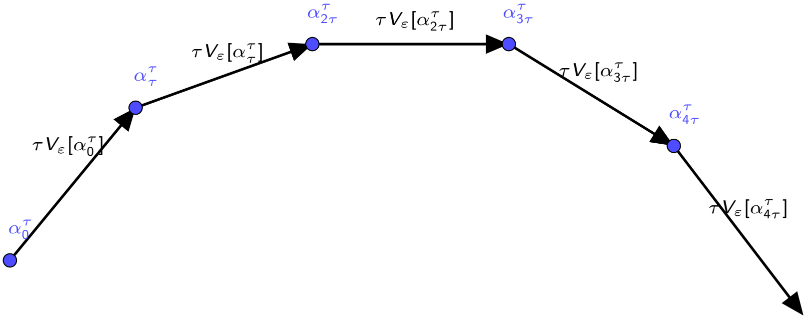

Figure: Euler scheme (ρ_t^τ)

Proposition (linear convergence)

If $\rho_0 \in \mathcal{P}(B_R)$, there is a unique weak solution (ρ_t) to (SG_ε) and (ρ_t^τ) weakly $*$ converges to ρ_t. Moreover

$$W_2(\rho_t^\tau, \rho_T) \leq e^{Ct} \frac{M}{C} \tau \tag{2}$$

Proposition (linear convergence)

If $\rho_0 \in \mathcal{P}(B_R)$, there is a unique weak solution (ρ_t) to (SG_ε) and (ρ_t^τ) weakly $*$ converges to ρ_t. Moreover

$$W_2(\rho_t^\tau, \rho_T) \leq e^{Ct}\frac{M}{C}\tau \tag{2}$$

One-sided lipschitz condition :

- $\rho_0, \beta_0 \in \mathcal{P}(B_R)$
- ρ_t solving $\partial_t \rho_t = -\nabla \cdot (\rho_t V_\varepsilon[\rho_t])$
- β_t one step Euler $\beta_t := (id + tV_\varepsilon[\beta_0])_\# \beta_0$:

Proposition (linear convergence)

If $\rho_0 \in \mathcal{P}(B_R)$, there is a unique weak solution (ρ_t) to (SG_ε) and (ρ_t^τ) weakly $*$ converges to ρ_t. Moreover

$$W_2(\rho_t^\tau, \rho_T) \leq e^{Ct}\frac{M}{C}\tau \tag{2}$$

One-sided lipschitz condition :

- $\rho_0, \beta_0 \in \mathcal{P}(B_R)$
- ρ_t solving $\partial_t \rho_t = -\nabla \cdot (\rho_t V_\varepsilon[\rho_t])$
- β_t one step Euler $\beta_t := (id + tV_\varepsilon[\beta_0])_{\#}\beta_0$:

Then

$$W_2(\rho_t, \beta_t) \leq (1 + tC)W_2(\rho_0, \beta_0) + Mt^2$$