OPTIMAL TRANSPORTATION SOLVERS

AND

VARIATIONAL RELAXATIONS OF EULER EQUATIONS

Jean-David Benamou (MOKAPLAN - EPC INRIA-CEREMADE)

`https://team.inria.fr/mokaplan/`

Contents

Euler Equations (1755)

On space domain $D \subseteq \mathbb{R}^d$ and time interval $[0, 1]$

$$(E) \begin{cases} \partial_t V + (V \cdot \nabla) V = -\nabla p \\ div(V) = 0 \\ V \cdot \nu = 0 \qquad on \quad \partial D \\ V(0, x) = V_0(x) \end{cases}$$

Global in time weak solution open for $d = 3$
(see survey `http://cvgmt.sns.it/paper/1714/`
Danieri-Figalli)

Lagrangian formulation, Arnold's geodesics (1966)

- Follow fluid particles $t \to G(t, X_0) \in D$

$$\begin{cases} \dot{G}(t,X) = V(t, G(t,X_0)) \\ G(0,X_0) = X_0 \qquad X_0 \in D \end{cases}$$

- From (E) we get

$$G(t,\cdot) \in \mathbb{S}diff(D) := \{S : D \to D;\ diffeomorphism\, s.t.\, det(\nabla S) = 1\}$$

and $\ddot{G}(t, X_0) = -\nabla p(t, G(t, X_0))$.

$$G = Arg \inf_{\begin{cases} S \in H^1([0,1], \mathbb{S}diff(D)) \\ S(0,.) = Id, \quad S(1,.) = G(1,.) \end{cases}} \int_{[0,1]\times X} ||\dot{S}(t, X_0)||^2 \, dX_0 \, dt$$

The Boundary value problem, relaxation

$$G = Arg \inf_{\begin{cases} S \in H^1([0,1], \mathbb{S}diff(D)) \\ S(0,.) = Id, \quad S(1,.) = S^* \end{cases}} \frac{1}{2} \int_{[0,1]\times X} ||\dot{S}(t, X_0)||^2 \, dX_0 dt$$

- Pbm : Lack of completeness of $\mathbb{S}diff(D) \subset L^2(D) \to$ relaxation.
- Set of measure preserving mapping $D \to D$ is complete in L^2.

$$\mathbb{S}(D) := \{S \in L^2(D, D);\ S_\# \mathcal{L}_D = \mathcal{L}_D\}$$

$\mathcal{L}_D$: Lebesgue measure on D
($\mathcal{L}_D(A) = \mathcal{L}_D(S^{-1}(A))$ For all measurable $A \subset D$)

Computations with permutations (Brenier 89)

- Discretize D as the union of N identical cubes, with centers $\{X_i\}_{i=1,N}$.
- $\Sigma_N :=$ Permutations of $\{1,2,..,N\}$ "dense" in $\mathbb{S}$ as $N \to \infty$.
- Piecewise linear time discretization $t_m,\ m = 0,..,M$ yields

$$\inf_{\left\{\begin{array}{c}\sigma_1,...\sigma_{M-1} \in (\Sigma_N)^{M-1} \\ \sigma_0 = Id_N \quad \sigma_M = \sigma^* \in \Sigma_N\end{array}\right.} \sum_{m=2,M} \sum_{i=1,N} \frac{M}{2} \left\|X_{\sigma_m(i)} - X_{\sigma_{m-1}(i)}\right\|^2$$

Computations with permutations (Brenier 89)

- Gauss-Seidel resolution approach (kth iteration)

$$\sigma_m^k = Arg\min_{\sigma\in\Sigma_N} M\sum_{i=1,N}||X_{\sigma_{m-1}^k(i)} - X_{\sigma(i)}||^2 + ||X_{\sigma(i)} - X_{\sigma_{m+1}^{k-1}(i)}||^2$$
$$= Arg\min_{\sigma\in\Sigma_N} \sum_{i=1,N}||X_{\sigma(i)} - \frac{X_{\sigma_{m-1}^k(i)} + X_{\sigma_{m+1}^{k-1}(i)}}{2}||^2$$

- Set

$$c_{ij} = ||X_j - \frac{X_{\sigma_{m-1}^k(i)} + X_{\sigma_{m+1}^{k-1}(i)}}{2}||^2$$

  It is a Linear Assigment problem ($\min_\sigma \sum_i, c_{i\,\sigma(i)}$).
  Each GS iteration costs $O(N^3)$ except in 1D $O(N\,LogN)$ (Sort).

- 1D tests from Brenier 2008.

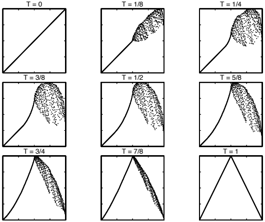


FIG. 1: APPROXIMATE GEODESIC FOR MAP 1

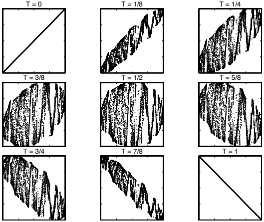


FIG. 3: APPROXIMATE GEODESIC FOR MAP 3

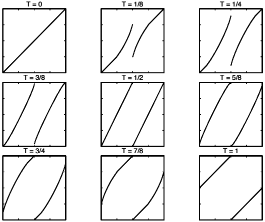


FIG. 2: APPROXIMATE GEODESIC FOR MAP 2

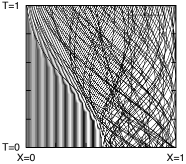


FIG. 4: TRAJECTORIES FOR MAP 1

Relaxation : generalized incompressible flows Brenier (89)

- Set $\Omega(D) = C([0,1];D)$ and e_t the evaluation map at time t, i.e. $e_t(\omega) = \omega(t)$ for $\omega \in \Omega(D)$

$$(P_\gamma) \quad \inf \left\{ \begin{array}{l} \gamma \in \mathcal{P}(\Omega(D)) \text{ s. t.} \\ (e_t)_\# \gamma = \mathcal{L}_D \text{ for all } t \\ (e_0,\, e_1)_\# \gamma = (Id,\, S^*)_\# \mathcal{L}_D \end{array} \right. \int_{\Omega(D)} \int_{[0,1]} \| \frac{1}{2} \dot{\omega}(t) \|^2 \, dt \, d\gamma(\omega)$$

- $(e_t)_\# \gamma(A) = d\gamma(\{\omega \in \mathcal{P}(\Omega(D)),\, \omega(t) \in A)\}$ forall $A \subset D$.
- $(e_0,\, e_t)_\# \gamma(A,B) = d\gamma(\{\omega \in \mathcal{P}(\Omega(D)),\, \omega(0) \in A \text{ and } \omega(t) \in B\}$ forall $(A,B) \subset D \times D$.
- $(Id,\, S^*)_\# \mathcal{L}_D(A,B) = \mathcal{L}_D(\{X \in A, \text{ s.t. } S^*(X) \in B\})$

Relaxation : generalized incompressible flows

- A flow $t \to G(t)$ in $\mathbb{S}\mathit{diff}(D)$ induces a generalized incompressible flow

$$\gamma_G \ : \ \gamma_G = (\Phi_G)_\# \mathcal{L}_D$$

where $\Phi_G : D \to \Omega(D)$, $\Phi_G(X_0) = G(.,X_0)$.

- The Kinetic energies "actions" coincide and

$$(e_0, e_1)_\# \gamma_G = (Id,\ G(1,.))_\# \mathcal{L}_D$$

- Reciprocally γ is a "deterministic" generalized incompressible flows at time t if there exist t a map $G_t \in \mathbb{S}$ such that

$$(e_0, e_t)_\# \gamma = (Id,\ G_t)_\# \mathcal{L}_D$$

Generalized incompressible flows : properties

- $\mathbb{S}diff$ dense in $\mathbb{S}$ for $d \geq 3$ but not for $d = 2$ (Shnirelman 85) : pathological examples with no minimizers in $\mathbb{S}diff$ for (P_S).
- Minimizers exists for (P_γ) and there is a unique pressure Lagrange multiplier of the incompressibility constraint.
- Consistency (Brenier 89) : given an admissible $\gamma \in \mathcal{P}(\Omega(D))$ and pressure p with correct initial-final coupling constraint and such that

$$\ddot{\omega}(t) = -\nabla p(t, \omega(t)), \quad \gamma - a.e.$$

$$\sup_{(t,X)\in[0,T]\times D} \nabla^2_X p(t,X) \leq \frac{\pi^2}{T^2} Id$$

  then γ solves (P_γ).
- If $< \leftarrow \leq$ above, γ is unique and deterministic.

"Free" Particle discretization of $\mathbb{S}$

Q. Merigot- J.-M. Mirebeau

https://hal.archives-ouvertes.fr/hal-01152168

- Define a flow of empirical measures $\nu_X(t_m) = \frac{1}{N}\sum_{i=1,..N}\delta_{X_i^m}$
- Instead of the "taquin" game, relax the measure preserving constraint by projecting on $\mathbb{S}$.
- Laguerre Diagrams

$$V_n(\{W_i\}_{i=1,..,N}) = \{X \in D,\ s.t.\ ||X - X_n||^2 + W_n \leq ||X - X_i||^2 + W_i,\ \ \forall i \neq k\}$$

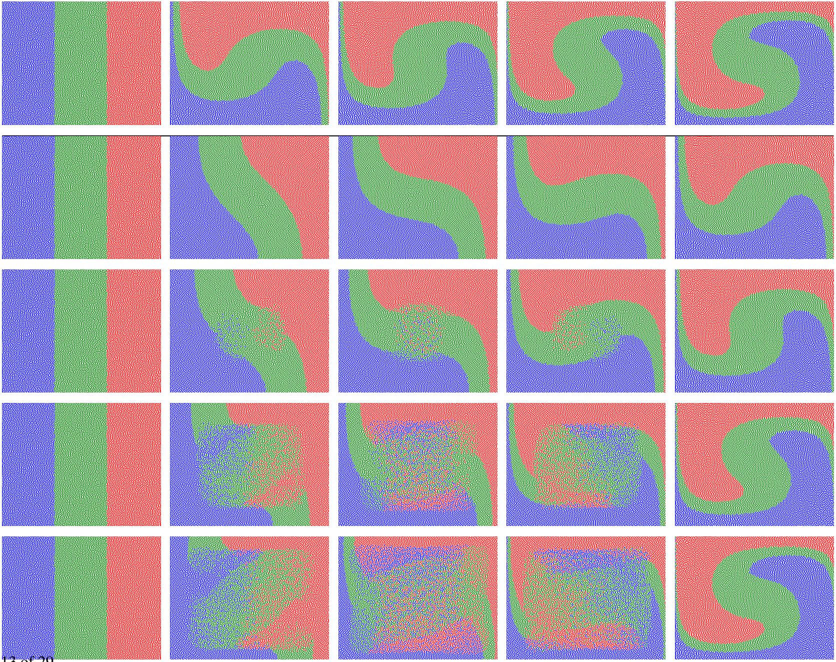
Like Voronoi D. but finding weights W such that $\mathcal{L}_D(V_n) = \frac{1}{N}$ amounts to solve the Quadratic Optimal Transport problem $W^2(\nu_X(t), \mathcal{L}_D) = \sum_n \int_{V_n} ||x - X_n||^2 dx.$

"Free" Particle discretization of $\mathbb{S}(D)$

- Solve

$$(P_{Mr[b,g]}) \begin{cases} \inf_{\{X_i^m\}} & \frac{M}{2}\sum_{m=2,M}\sum_{i=1,N}\|X_i^m - X_i^{m-1}\|^2 + \\ & \lambda\sum_{k=2,M-1} W^2(\nu_X(t_m), \mathcal{L}_D) + \\ & \lambda\left(\|X_i^1 - Id_i\|^2 + \|X_i^M - S_i^*\|^2\right) \end{cases}$$

- Generalised flow measure (γ) formulation possible. Convergence results, Fast Quasi Newton algorithm $O(M\,N)$.
- Non convex. Smart initialisation needed.
- Test (Brenier Roesch) Beltrami Stationary flow on $D = [0,1]^2$
$V(x,y) = \{-\cos(\pi x)\sin(\pi y), \sin(\pi x)\cos(\pi y)\}$
$p(x,y) = \frac{1}{4}(\sin^2(\pi x) + \sin^2(\pi y))\ S^\star = G(T,.)$.



Piecewise linear time discretization of (P_γ)

$$(P_\gamma) \qquad \inf_{\begin{cases} \gamma \in \mathcal{P}(\Omega(D)) \text{ s. t.} \\ (e_t)_\# \gamma = \mathcal{L}_D \text{ for all } t \\ (e_0, e_1)_\# \gamma = (Id \times S^*)_\# \mathcal{L}_D \end{cases}} \int_{\Omega(D)} \int_{[0,1]} \frac{1}{2} ||\dot{\omega}(t)||^2 \, dt \, d\gamma(\omega)$$

- $X_m = \omega(t_m)$ and

$$c(X_1, X_2, .., X_M) = \frac{M}{2} \sum_{m=2,M} ||X_m - X_{m-1}||^2 \quad \left[+\lambda \, ||X_M - S^*(X_1)||^2\right]$$

$$\inf_{\begin{cases} \gamma_M \in \mathcal{P}(D^M) \text{ s. t.} \\ \pi_{m\#} \gamma_M = \mathcal{L}_D \text{ for all } m \\ [(\pi_0, \pi_M)_\# \gamma_M = (Id \times S^*)_\# \mathcal{L}_D] \end{cases}} \int_{D^M} c(X_1, X_2, .., X_M) \, d\gamma_M(X_1, X_2, .., X_M)$$

Space discretization X_{i_m} for $i = 1, .., N$ at time t_m.

$$c_{i_1,i_2,..,i_M} = \frac{M}{2} \sum_{m=2,M} \|X_{i_m} - X_{i_{m-1}}\|^2 \quad \left[+\lambda \|X_{i_M} - S^*(X_{i_1})\|^2\right]$$

- t_m-marginal : $\gamma^1_{i_m} = \sum_{i_1,..,i_{m-1},\not{i_m},i_{m+1},..,i_M=1,N} \gamma_{i_1,i_2,..,i_M}$
- Probability of transition from t_1 to t_m :

$$\gamma^2_{i_1,i_m} = \sum_{\not{i_1},i_2,..,i_{m-1},\not{i_m},i_{m+1},..,i_M=1,N} \gamma_{i_1,i_2,..,i_M}$$

$$\begin{cases} \inf\limits_{\{\gamma_{i_1,i_2,..,i_M}\} \in (\mathbb{R}_+)^{N^M} \text{ s. t.}} \sum\limits_{i_1,i_2,..,i_M=1,N} c_{i_1,i_2,..,i_M} \gamma_{i_1,i_2,..,i_M} \\ \gamma^1_{i_m} = \frac{1}{N} \text{ for all } i_m \\ \left[\gamma^2_{i_1,i_M} = \delta_{i_M,\sigma_{S^*}(i_1)} \text{ for all } (i_1, i_M)\right] \end{cases}$$

Entropic regularization of OT (see Leonard Survey `https://arxiv.org/abs/1308.0215` for the connection with Schrödinger problem)

$$\min_{\begin{cases} \gamma_{ij}^{\epsilon} \geq 0 \\ \sum_j \gamma_{ij}^{\epsilon} = \mu_i \ , \sum_i \gamma_{ij}^{\epsilon} = \nu_j \end{cases}} \sum_{ij} \gamma_{ij}^{\epsilon} c_{ij} + \epsilon \gamma_{ij}^{\epsilon} (log \gamma_{ij}^{\epsilon} - 1)$$

Set $\overline{\gamma}_{ij}^{\epsilon} = e^{-\frac{c_{ij}}{\epsilon}}$

$$\min_{\begin{cases} \gamma_{ij}^{\epsilon} \\ \sum_j \gamma_{ij}^{\epsilon} = \mu_i \ , \sum_i \gamma_{ij}^{\epsilon} = \nu_j \end{cases}} \epsilon \sum_{ij} KL(\gamma_{ij}^{\epsilon} | \overline{\gamma}_{ij}^{\epsilon})$$

$$KL(f|g) = f\,(\log(\frac{f}{g}) - 1)$$

Iterative Proportional Fitting Procedure Sinkhorn (67) Ruschendorf (95) Galichon (09) Cuturi (13) ...

$$\min_{\{\gamma^\epsilon_{ij}\}} \max_{\{\phi^\epsilon_i, \psi^\epsilon_j\}} \sum_{ij} \psi^\epsilon_j \, \nu_j + \phi^\epsilon_i \, \mu_i + \gamma^\epsilon_{ij} \left(c_{ij} - \psi^\epsilon_j - \phi^\epsilon_i + \epsilon \left(\log \gamma^\epsilon_{ij} - 1 \right) \right)$$

- $\gamma^{\star,\epsilon}_{ij} = a^\epsilon_i \, b^\epsilon_j \, \overline{\gamma}^\epsilon_{ij}$ where $a^\epsilon_i = e^{\frac{\phi^\epsilon_i}{\epsilon}}$ and $b^\epsilon_j = e^{\frac{\psi^\epsilon_j}{\epsilon}}$.
- Margin constraints give $a^\epsilon_i = \dfrac{\mu_i}{\sum_j \overline{\gamma}^\epsilon_{ij} \, b^\epsilon_j}$ and $b^\epsilon_j = \dfrac{\nu_j}{\sum_i \overline{\gamma}^\epsilon_{ij} \, a^\epsilon_i}$.
- IPFP is the relaxation: $b^{\epsilon,0} = 1$

$$a_i^{\epsilon,k+\frac{1}{2}} = \frac{\mu_i}{\sum_j \overline{\gamma}^\epsilon_{ij} \, b_j^{\epsilon,k}} \qquad b_j^{\epsilon,k+1} = \frac{\nu_j}{\sum_i \overline{\gamma}^\epsilon_{ij} \, a_i^{\epsilon,k+\frac{1}{2}}}.$$

IPFP = Bregman iterative projection for KL divergence (67)

- The Entropic regularization of OT is

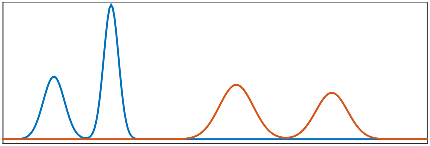
$$P^{KL}_{\gamma\in C}(\overline{\gamma}) = Argmin_{\gamma\in C} \sum_{ij} KL(\gamma^{\epsilon}_{ij}|\overline{\gamma}^{\epsilon}_{ij})$$

  where $C = C_\mu \cap C_\nu = \{\gamma, \sum_j \gamma^{\epsilon}_{ij} = \mu_i\} \cap C_\mu = \{\gamma, \sum_i \gamma^{\epsilon}_{ij} = \nu_j\}$ is the intersection of linear (Convex) subspaces ..

- Bregman iterative alternate projection converge (Bauske Lewis, ...) $\gamma^0 = \overline{\gamma}$ then $\forall k\ \gamma^{k+1} = P^{KL}_{\gamma\in C_\mu}(\gamma^k)$ and $\gamma^{k+2} = P^{KL}_{\gamma\in C_\nu}(\gamma^{k+1})$.

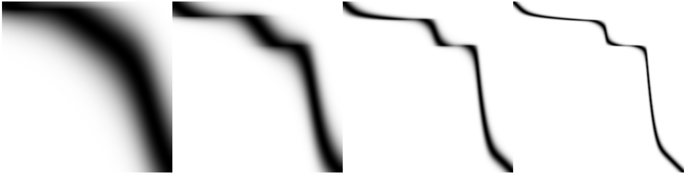
- Projections are explicit equivalent to IPFP with $\gamma^k_{ij} = a_i^{\epsilon,k-\frac{1}{2}}\overline{\gamma}^{\epsilon}_{ij}b_j^{\epsilon,k}$ $\gamma^{k+1}_{ij} = a_i^{\epsilon,k+\frac{1}{2}}\overline{\gamma}^{\epsilon}_{ij}b_j^{\epsilon,k}$

Data :



Iterations :

$\epsilon \rightarrow 0$:



Remarks on IPFP

- Convergence with ϵ (Cominetti San MArtin (94) in the OT context see Carlier et all `http://arxiv.org/abs/1512.02783`.
- Costs $N^2 \times \#iterations$, store N^2 matrices
- $\#$ iterations increase with $\frac{1}{\epsilon}$.
- Easy implementation (does not depend on discretisation or dimension of the problem).
- Applies to general cost matrix c_{ij}.
- Generalizes to many variants of OT including multimarginal check B. et all `http://arxiv.org/abs/1412.5154`
- If the plan concentrates, heuristics to refine and restrict its support see Schmitzer `http://arxiv.org/abs/1510.05466` and Oberman.

IPFP for the Multi-Marginal problem

- $KL(\gamma|\overline{\gamma}) = \sum_{ijk} \gamma_{ijk} \log \frac{\gamma_{ijk}}{\overline{\gamma}_{ijk}}$ with $\overline{\gamma}_{ijk} = e^{-\frac{c_{ijk}}{\epsilon}}$ and $\mathcal{C} = \bigcap_{i=1}^{3} \mathcal{C}_i$ (i.e. $\mathcal{C}_i = \{\gamma \quad | \quad \sum_{j,k} \gamma_{ijk} = \rho_i\}$).
- The optimal plan $\gamma^\star$ becomes $\gamma^\star_{ijk} = a_i b_j c_k \overline{\gamma}_{ijk}$

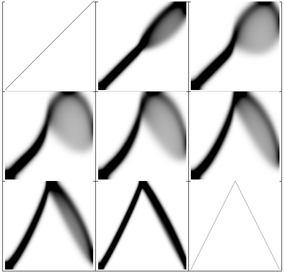
$$b_j = \frac{\rho_j}{\sum_{ik} a_i c_k \overline{\gamma}_{ijk}} \qquad b_j^{n+1} = \frac{\rho_j}{\sum_{ik} a_i^n c_k^n \overline{\gamma}_{ijk}}$$

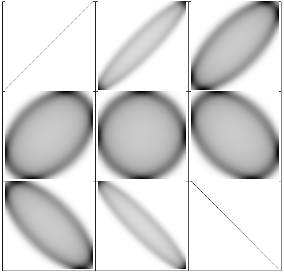
$$c_k = \frac{\rho_k}{\sum_{ij} a_i b_j \overline{\gamma}_{ijk}} \qquad {c_k}^{n+1} = \frac{\rho_k}{\sum_{ij} a_i^n b_j^{n+1} \overline{\gamma}_{ijk}}$$

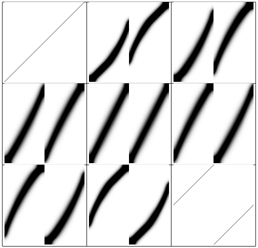
$$a_i = \frac{\rho_i}{\sum_{jk} b_j c_k \overline{\gamma}_{ijk}} \qquad a_i^{n+1} = \frac{\rho_i}{\sum_{jk} b_j^{n+1} c_k^{n+1} \overline{\gamma}_{ijk}}$$

IPFP for Euler

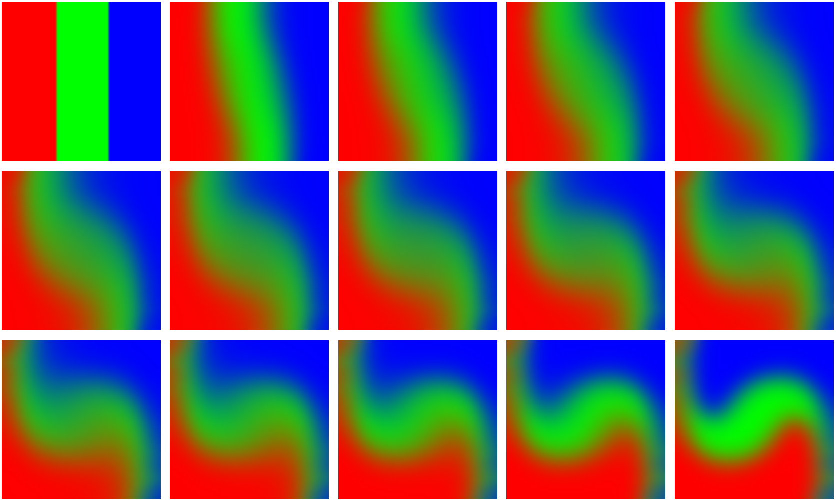
- IPFP algebra amounts to $u_{i_m}^{m,(k)} = \frac{1/N}{\sum_{i_1,..,i_{m-1},\not{i_m},i_{m+1},..,i_M} u_{i_1}^{1,(k)} u_{i_2}^{2,(k)} \dots u_{i_{m-1}}^{m-1,(k)} u_{i_{m+1}}^{m+1,(k-1)} \dots u_{i_{M-1}}^{M-1,(k-1)} u_{i_M}^{M,(k-1)} \overline{\gamma}_{i_1,\dots,i_M}}$
- Cost metric is separable (also with dimensions) $\overline{\gamma}_{i_1,\dots,i_M} = \prod_{m=1}^{M-1} \xi^0_{i_m\, i_{m+1}}\, \xi^1_{i_M\, i_1}$
- $\xi^0_{\alpha\,\beta} = e^{-\frac{||x_\alpha - x_\beta||^2}{\epsilon}}$ and $\xi^1_{\alpha\,\beta} = e^{-\lambda\,\frac{||x_\alpha - S^*(x_\beta)||^2}{\epsilon}}$.
- $\rightarrow$ store N^2 matrices
- $\rightarrow$ one iteration costs $O((M\,N^{1.5})$



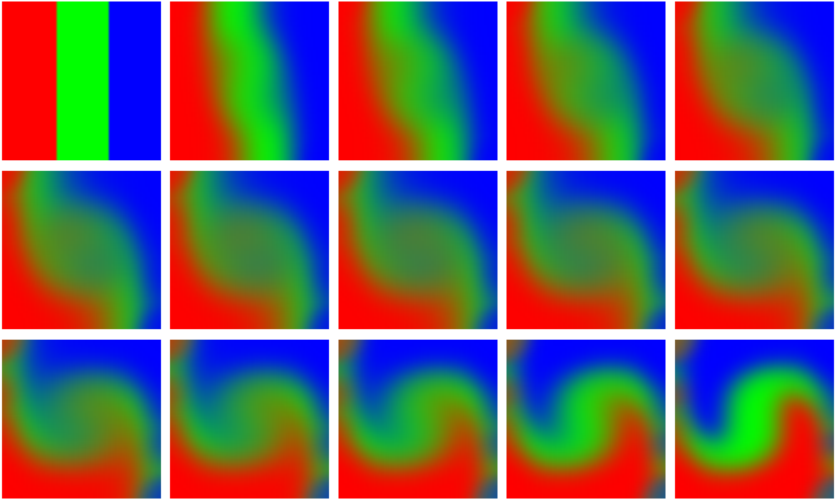




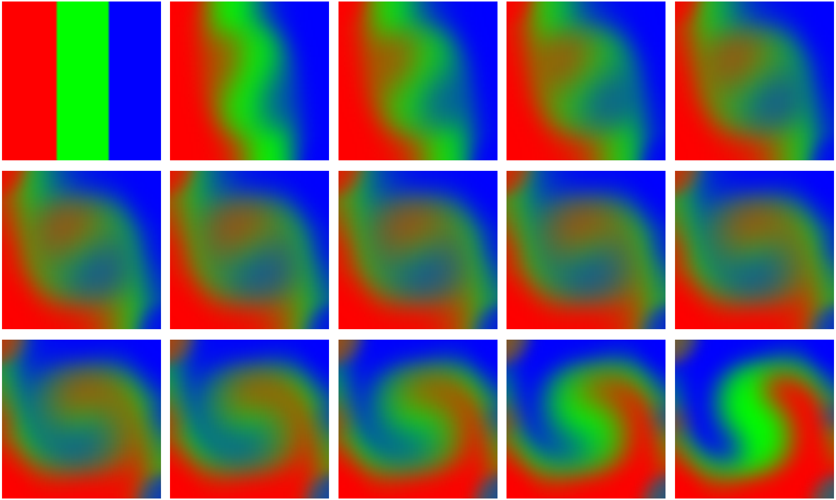
Test Beltrami $T = 0.9$



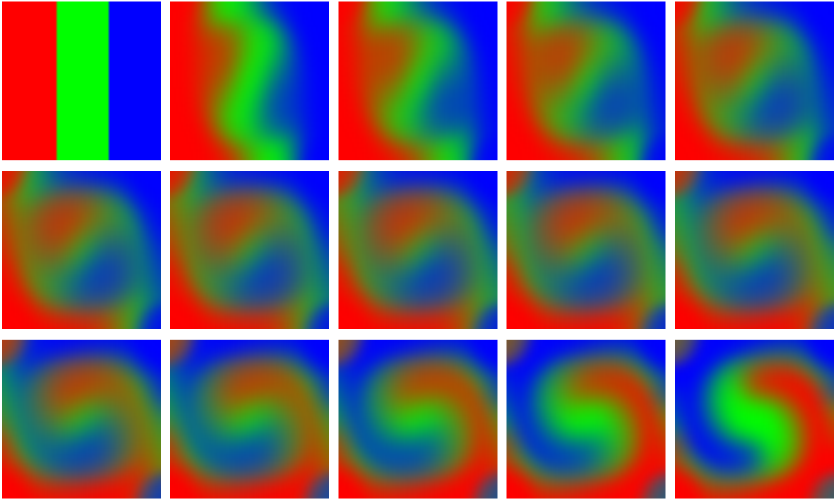
Test Beltrami $T = 1.1$



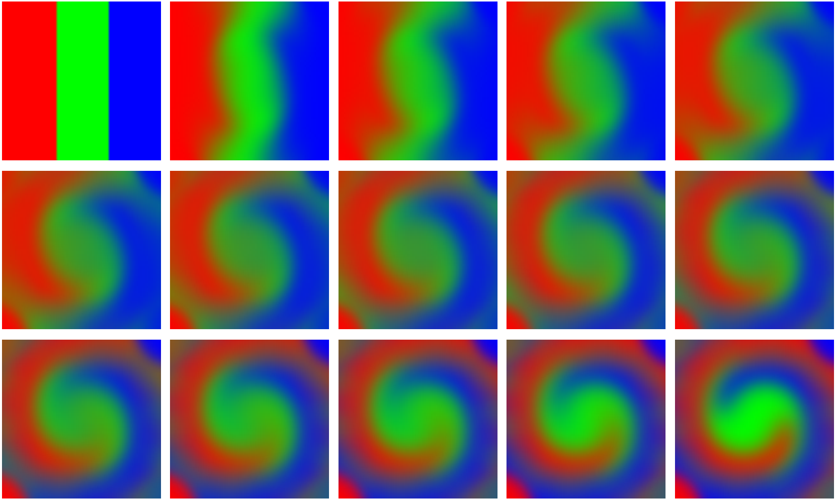
Test Beltrami $T = 1.3$

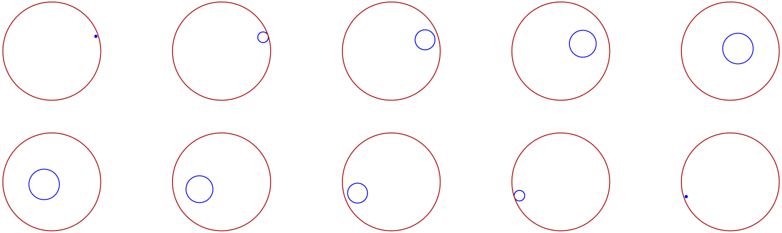


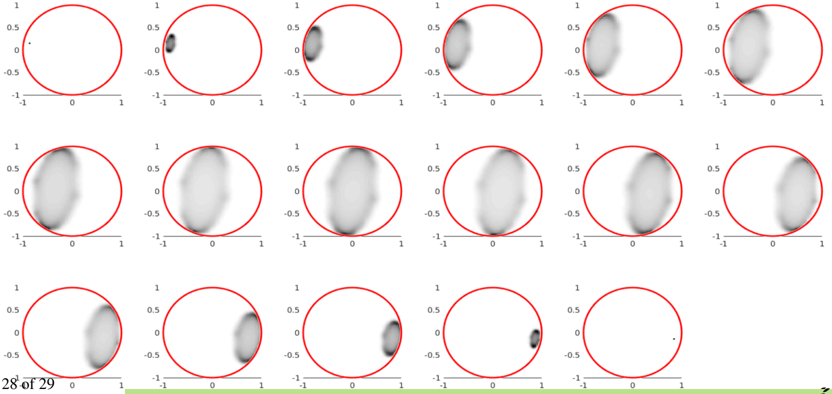
Test Beltrami $T = 1.5$



Test Beltrami $T = \pi$






1
0.5
0
-0.5
-1
-1
0
1

THANK YOU

Register for `https://project.inria.fr/brenier60/` if you are interested.