

La mécanique des fluides, un problème de transport ?

Thomas Jacumin

17 novembre 2022

Table des matières

1	L'équation de la continuité	2
1.1	L'équation de la continuité	2
1.2	Formulation faible	2
1.3	Le point de vue Lagrangien	3
2	Le lien avec le transport	4
2.1	L'équation de la continuité pour des mesures	4
2.2	Formule de Benamou-Brenier	5
3	La méthode du Lagrangien (Augmenté)	7
3.1	Le Lagrangien (Augmenté)	7
3.2	Un algorithme de Uzawa	8
4	Le Lagrangien augmenté pour la formule de Benamou-Brenier	8
5	Algorithme d'Uzawa	10
5.1	Étape 1	10
5.2	Étape 2	11
5.3	Étape 3	11
5.4	Condition d'arrêt	11
6	Résultats numériques	11
	Références	12

On a vu que, sous les bonnes hypothèses et en utilisant le théorème de Brenier (théorème 2.2), l'application de transport T optimal pouvait s'écrire comme un flot de gradient i.e.

$$T = \nabla \Psi,$$

avec Ψ convexe, s.c.i.

$$\liminf_{x \rightarrow x_0} f(x) \leq f(x_0), \quad \forall x_0,$$

et propre i.e.

$$\Psi > -\infty, \quad \Psi \not\equiv +\infty.$$

De plus, Ψ est une solution, au sens adéquat, de l'équation de Monge-Ampère,

$$\det(H\Psi(x)) \nu(\nabla\Psi(x)) = \mu(x).$$

Cette équation est non-linéaire et n'avait pas été trop étudiée pour sa résolution numérique jusque dans les années 2000. En 2000, Benamou et Brenier [1] ont réécrit ce problème de transport optimal comme un problème de mécanique des fluides. L'avantage de cette formulation est d'être linéaire et facilement résolvable numériquement.

1 L'équation de la continuité

Soit D un ouvert régulier et borné de \mathbb{R}^d . Soient deux fonctions μ et ν de \mathbb{R}^d dans \mathbb{R}_+ à support dans D et telle que

$$\int_D \mu \, dx = \int_D \nu \, dx = 1.$$

1.1 L'équation de la continuité

On considère le problème suivant :

Problème 1.1 (Équation de la continuité (au sens fort)). *Trouver (ρ, v) tel que, $\rho \geq 0$ et*

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = 0, & \text{dans } [0, 1] \times D, \\ \rho(0, \cdot) = \mu \text{ et } \rho(1, \cdot) = \nu, & \text{dans } D, \\ v \cdot n = 0, & \text{sur } [0, 1] \times \partial D. \end{cases}$$

Remarque 1. On a pour A ouvert régulier inclut dans D ,

$$\frac{d}{dt} \int_A \rho(t, x) \, dx = - \int_A \operatorname{div}(\rho(t, x) v(t, x)) \, dx = \int_A \rho(t, x) \nabla \cdot v(t, x) - \int_{\partial A} (\rho v) \cdot n \, d\sigma = - \int_{\partial A} (\rho v) \cdot n \, d\sigma,$$

avec n la normale unitaire extérieure de ∂D . Ainsi, la variation de la masse dans A est égale à ce qui rentre dans A moins ce qui sort de A .

Remarque 2. Pour tout t on peut définir l'énergie cinétique des particules par

$$E(t) := \int_D \rho(t, x) |v(t, x)|^2 \, dx.$$

On peut aussi définir l'action du champ de vélocité v par

$$A[\rho, v] := \int_0^1 E(t) \, dt.$$

Il s'agit de l'énergie totale nécessaire pour déplacer les particules le long de v .

Remarque 3. L'hypothèse $v \cdot n = 0$ sur ∂D permet d'assurer que le flot v reste à l'intérieur du domaine D :

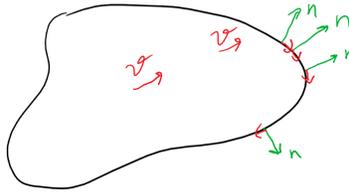


FIGURE 1 – Condition $v \cdot n = 0$ sur ∂D .

On a la conservation de la masse : pour tout t dans $[0, 1]$,

$$\frac{d}{dt} \int_D \rho(t, x) \, dx = - \int_{\partial D} (\rho v) \cdot n \, d\sigma = 0.$$

1.2 Formulation faible

Pour que l'équation de la continuité au sens fort est un sens, il faut imposer des hypothèses fortes sur la paire de solution (ρ, v) . Afin de s'en débarrasser, on introduit la notion de problème faible de la manière suivante : soit

$\phi \in C_c^1([0, 1] \times \bar{D}, \mathbb{R})$, alors,

$$\begin{aligned}
& \int_0^1 \int_D (\partial_t \rho + \operatorname{div}(\rho v)) \phi \, dx \, dt = 0 \\
\Leftrightarrow & \int_0^1 \int_D \partial_t \rho \phi \, dx \, dt + \int_0^1 \int_D \operatorname{div}(\rho v) \phi \, dx \, dt = 0 \\
\Leftrightarrow & \int_D \rho(1, \cdot) \phi(1, \cdot) \, dx - \int_D \rho(0, \cdot) \phi(0, \cdot) \, dx - \int_0^1 \int_D \rho \partial_t \phi \, dx \, dt + \int_0^1 \int_{\partial D} \underbrace{(\rho v) \cdot n}_{=0} \phi \, d\sigma \, dt \\
& \quad - \int_0^1 \int_D \rho v \cdot \operatorname{grad}(\phi) \, dx \, dt = 0 \\
\Leftrightarrow & \int_0^1 \int_D \rho \partial_t \phi \, dx \, dt + \int_0^1 \int_D \rho v \cdot \operatorname{grad}(\phi) \, dx \, dt = \int_D \rho(0, \cdot) \phi(0, \cdot) \, dx - \int_D \rho(1, \cdot) \phi(1, \cdot) \, dx \\
\Leftrightarrow & \int_0^1 \int_D \rho \partial_t \phi \, dx \, dt + \int_0^1 \int_D \rho v \cdot \operatorname{grad}(\phi) \, dx \, dt = \int_D \mu \phi(0, \cdot) \, dx - \int_D \nu \phi(1, \cdot) \, dx.
\end{aligned}$$

On peut donner l'équation de la continuité au sens faible :

Problème 1.2 (Équation de la continuité (au sens faible)). *Trouver (ρ, v) avec $\rho \geq 0$, $(\rho(t, \cdot) v(t, \cdot)) \in L^1(D)$ et $\int_0^1 \|\rho(t, \cdot) v(t, \cdot)\|_{L^1(D)} \, dt < +\infty$ tel que, pour tout $\phi \in C_c^1([0, 1] \times \bar{D}, \mathbb{R})$, on ait*

$$\int_0^1 \int_D \rho \partial_t \phi \, dx \, dt + \int_0^1 \int_D \rho v \cdot \operatorname{grad}(\phi) \, dx \, dt = \int_D \mu \phi(0, \cdot) \, dx - \int_D \nu \phi(1, \cdot) \, dx.$$

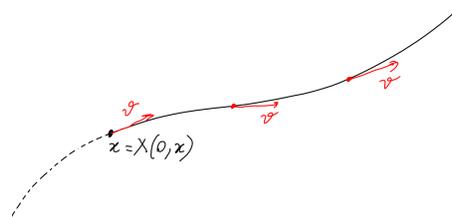
Maintenant, on peut faire le lien entre la formulation faible et la formulation forte au moyen de la proposition suivante [3] :

Proposition 1.1. *Soit (ρ, v) une solution du problème faible. Si ρ est Lipschitzienne en (t, x) et v est Lipschitzienne en x , alors (ρ, v) est aussi une solution du problème fort p.p. dans $[0, 1] \times \bar{D}$.*

1.3 Le point de vue Lagrangien

Le point de vue Eulérien consiste à observer les propriétés telles que la pression ou la vitesse au cours du temps à un endroit fixé de l'espace. Le point de vue Lagrangien est une façon de regarder le mouvement des fluides où l'observateur suit une parcelle de fluide individuelle lorsqu'elle se déplace dans l'espace et le temps. Plus précisément, on définit $X : [0, 1] \times D \rightarrow \mathbb{R}^d$ comme une solution (si elle existe) du problème de Cauchy suivant :

$$\begin{cases} X(0, x) = x, & \forall x \in D, \\ \partial_t X(t, x) = v(t, X(t, x)), & \forall (t, x) \in [0, 1] \times D. \end{cases} \quad (1)$$



On a le résultat d'existence d'une solution au problème de Cauchy grâce au théorème de Cauchy-Lipschitz :

Théorème 1.1 (Théorème de Cauchy-Lipschitz). *Si v est lipschitzienne en x uniformément en t i.e. il existe $L \in \mathbb{R}$ (indépendante de t et de x) telle que pour tout $(t, x, y) \in [0, 1] \times D \times D$,*

$$|v(t, x) - v(t, y)| \leq L|x - y|,$$

alors (1) admet une unique solution X et pour tout $t \in [0, 1]$, $X(t, \cdot)$ est lipschitzienne, inversible et son inverse, $(X(t, \cdot))^{-1}$, est aussi lipschitzienne.

Enfin, on a

Proposition 1.2. *Avec les mêmes hypothèses sur v que pour le théorème de Cauchy-Lipschitz, si on suppose de plus que $v \cdot n = 0$ sur ∂D alors le support de X , X solution de (1), est inclut dans D .*

2 Le lien avec le transport

Dans cette section, on va présenter et démontrer la formule de Benamou-Brenier [1].

2.1 L'équation de la continuité pour des mesures

Tout d'abord, on étend l'équation de la continuité au cas où ρ n'est plus une densité mais une mesure. Pour bien faire la différence entre densité et mesure, on ajoutera une barre au dessus du symbole pour indiquer qu'il s'agit d'une mesure. On dit que $(\bar{\rho}, v)$ est une solution faible lorsque,

$$\int_0^1 \int_D \partial_t \phi \, d\bar{\rho} \, dt + \int_0^1 \int_D v \cdot \text{grad}(\phi) \, d\bar{\rho} \, dt = \int_D \phi(0, \cdot) \, d\mu - \int_D \phi(1, \cdot) \, d\nu.$$

Remarque 4. On a pas unicité de la solution, en effet, si $v_1 = 0$, alors $\mu = \nu = \bar{\rho}(t, \cdot)$ et X_1 est solution et si v_2 décrit des cercles (voir figure 2) tel qu'on ait toujours $\mu = \nu$ alors, X_2 est aussi solution mais $X_1 \neq X_2$.

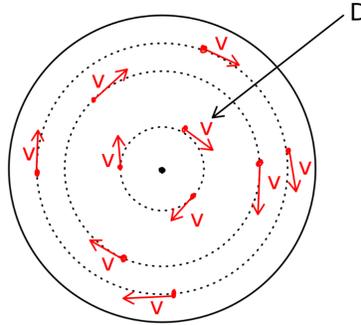


FIGURE 2 – Champs de vecteur v décrivant des cercles.

Remarque 5. Si $\bar{\rho}$ est absolument continue de densité ρ alors on retrouve la formulation faible de l'équation de la continuité présenté dans la section précédente.

On a la proposition suivante, démontrée dans [4], Théorème 5.34.

Théorème 2.1. *Soit $X : [0, 1] \times D \rightarrow D$ telle que pour tout $t \in [0, 1]$, $X(t, \cdot)$ est lipschitzienne, inversible et son inverse, $(X(t, \cdot))^{-1}$, est aussi lipschitzienne. Soit μ une mesure de probabilité sur D . On pose*

$$v(t, x) := \partial_t X \left(t, (X(t, \cdot))^{-1}(x) \right), \quad \forall (t, x) \in [0, 1] \times D,$$

et

$$\bar{\rho}(t, \cdot) := X(t, \cdot) \# \mu, \quad \forall t \in [0, 1].$$

Alors, $(\bar{\rho}, v)$ est solution de l'équation de la continuité au sens faible ayant comme densité initiale $\mu = \bar{\rho}(0, \cdot) = \text{Id} \# \mu$ et finale $\nu := \bar{\rho}(1, \cdot) = X(1, \cdot) \# \mu$.

Remarque 6. En utilisant la proposition 1.2, on a en particulier que $\bar{\rho}(t, \cdot)$ est a support dans D pour tout $t \in [0, 1]$.

2.2 Formule de Benamou-Brenier

On suppose que μ et ν sont des mesures de probabilité sur D . On commence par rappeler la définition de la distance de Wasserstein.

Définition 2.1 (Distance de Wasserstein).

$$W_2(\mu, \nu) := \inf_{\pi \in \text{Marg}(\mu, \nu)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |y - x|^2 d\pi(x, y) \right)^{1/2}.$$

Et maintenant on rappelle le théorème de Brenier.

Théorème 2.2 (Brenier). Soient $\mu \in \mathcal{P}_2^{ac}(\mathbb{R}^d)$ et $\nu \in \mathcal{P}_2(\mathbb{R}^d)$.

- Si $\pi \in \mathcal{O}_2(\mu, \nu)$, alors il existe $T \in \text{Tr}(\mu, \nu)$ telle que $\pi = (id, T)_{\#}\mu =: \pi_T$. De plus, il existe $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ s.c.i., convexe et propre telle que $T = \nabla\varphi$ μ -p.p.
- $|\mathcal{O}_2(\mu, \nu)| = 1$. On note l'unique élément π^* .
- Si $\pi^* = \pi_{T_1} = \pi_{T_2}$, alors $T_1 = T_2$ μ -p.p.

Remarque 7. On avait vu que la distance de Wasserstein faisait intervenir les plans de transport. Lorsque le théorème de Brenier s'applique, on peut réécrire la distance de Wasserstein en terme d'application de transport. Ainsi, l'infimum est atteint en $\pi^* = (id, T)_{\#}\mu$ et donc

$$W_2(\mu, \nu) = \inf_{T \in \text{Tr}(\mu, \nu)} \left(\int_{\mathbb{R}^d} |T(x) - x|^2 d\mu(x) \right)^{1/2}.$$

Il reste un dernier rappel avant d'énoncer la formule de Benamou-Brenier.

Proposition 2.1. Soient $\mu, \nu \in \mathcal{P}^{ac}(\mathbb{R}^d)$ à support compact. Soient $T \in \text{Tr}(\mu, \nu)$ et $S \in \text{Tr}(\nu, \mu)$ telles que $\pi_T \in \mathcal{O}_2(\mu, \nu)$ et $\pi_S \in \mathcal{O}_2(\nu, \mu)$, alors,

- $(S \circ T)(x) = x$, μ -p.p.,
- $(T \circ S)(y) = y$, ν -p.p.,

Maintenant, on pose $V(\mu, \nu)$ l'ensemble des $(\bar{\rho}, v)$ tels que :

- $t \mapsto \bar{\rho}(t, \cdot)$ est dans $C^0([0, 1], \mathcal{P}^{ac}(D))$,
- v est dans $L^2(\bar{\rho}(t, \cdot) dt)$,
- $(\bar{\rho}, v)$ est une solution de l'équation de la continuité au sens faible,

Théorème 2.3 (Théorème de Benamou-Brenier). Soient $\mu, \nu \in \mathcal{P}^{ac}(\mathbb{R}^d)$ à support dans D . Alors,

$$W_2(\mu, \nu)^2 = \inf_{(\bar{\rho}, v) \in V(\mu, \nu)} \int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt =: \inf_{(\bar{\rho}, v) \in V(\mu, \nu)} A[\bar{\rho}, v].$$

Démonstration. • On montre que $W_2(\mu, \nu)^2 \leq \inf_{(\bar{\rho}, v) \in V_{\text{sm}}(\mu, \nu)} \int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt$, avec $V_{\text{sm}}(\mu, \nu) := \{(\bar{\rho}, v) \in V(\mu, \nu) \mid v \in C^1([0, 1] \times \bar{D})\}$.

Soit $(\bar{\rho}, v) \in V_{\text{sm}}(\mu, \nu)$. On passe au point de vue Lagrangien en utilisant le problème de Cauchy (1). Puisque v est C^1 sur $[0, 1] \times \bar{D}$ elle est Lipschitzienne en x uniformément en t donc, en utilisant le théorème de Cauchy-Lipschitz (théorème 1.1), on a l'existence et l'unicité de X et pour tout $t \in [0, 1]$, $X(t, \cdot)$ est lipschitzienne, inversible et son inverse, $(X(t, \cdot))^{-1}$, est aussi lipschitzienne. De plus, avec le théorème 2.1, on sait que $\bar{\rho}(t, \cdot) = X(t, \cdot)_{\#}\mu$ et que $(\bar{\rho}, v)$ est une solution faible de l'équation de la continuité. Ainsi,

$$\int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt = \int_0^1 \int_D |v(t, X(t, x))|^2 d\mu(x) dt = \int_0^1 \int_D |\partial_t X(t, x)|^2 d\mu(x) dt.$$

En utilisant Fubini et l'inégalité de Jensen, on obtient,

$$\int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt \geq \int_D \left| \int_0^1 \partial_t X(t, x) dt \right|^2 d\mu(x) = \int_D |X(1, x) - x|^2 d\mu(x).$$

Mais puisque $X(1, \cdot) \# \mu = \bar{\rho}(1, \cdot) = \nu$, alors $X(1, \cdot)$ est une application de transport entre μ et ν et donc

$$\int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt \geq W_2(\mu, \nu)^2.$$

- On montre que $W_2(\mu, \nu)^2 \leq \inf_{(\bar{\rho}, v) \in V(\mu, \nu)} \int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt$.

On utilise un argument d'approximation pour ramener le problème de minimisation au cas des champs de vélocité v réguliers et ainsi pouvoir utiliser le résultat de l'étape 1 pour conclure (voir [4] Théorème 8.1 pour plus de détails).

- On montre qu'il existe $(\bar{\rho}_p, v_p) \in V(\mu, \nu)$ tel que $W_2(\mu, \nu)^2 = \int_0^1 \int_D |v_p(t, x)|^2 d\bar{\rho}_p(t, x) dt$.

Soit $\nabla\psi$ réalise le transfert de μ vers ν et, avec le théorème de Brenier, $\nabla\psi$ est optimale. On définit (coordonnées Lagrangiennes) $X : [0, 1] \times D \rightarrow D$, tel que,

$$X(t, x) = (1-t)x + t\nabla\psi(x) = \nabla \left(\underbrace{(1-t)\frac{|x|^2}{2}}_{\text{strictement convexe}} + \underbrace{t\psi(x)}_{\text{convexe}} \right) =: \nabla\Psi(t, x).$$

Ainsi, $X(0, x) = x$ et $X(1, x) = \nabla\psi(x)$ et on a

$$(1-t)\langle x-y, x-y \rangle \leq (1-t)\langle x-y, x-y \rangle + t \underbrace{\langle \nabla\psi(x) - \nabla\psi(y), x-y \rangle}_{\geq 0 \text{ car } \psi \text{ convexe}} = \langle \nabla\Psi(t, x) - \nabla\Psi(t, y), x-y \rangle.$$

D'où pour $0 \leq t < 1$, avec Cauchy-Schwarz

$$|x-y|^2 \leq \frac{1}{(1-t)} \langle \nabla\Psi(t, x) - \nabla\Psi(t, y), x-y \rangle \leq C_t |\nabla\Psi(t, x) - \nabla\Psi(t, y)| |x-y|.$$

Alors, on a que

$$\mathcal{L}^d(X^{-1}(t, A)) \leq C(C_t)^d \mathcal{L}^d(A),$$

en utilisant la définition de la mesure de Hausdorff, $\text{diam}(X(t, A)^{-1}) \leq C_t \text{diam}(A)$ et $\mathcal{L}^d(A) = C \mathcal{H}^d(A)$. Ainsi,

$$\bar{\rho}_p(t, \cdot) := X(t, \cdot) \# \bar{\mu},$$

est absolument continue par rapport à la mesure de Lebesgue. En effet, Soit A tel que $\mathcal{L}^d(A) = 0$. Comme $0 \leq \mathcal{L}^d(X^{-1}(t, A)) \leq C(C_t)^d \mathcal{L}^d(A) = 0$, alors $\mathcal{L}^d(X^{-1}(t, A)) = 0$. Comme μ est absolument continue, on a également que $\bar{\rho}_p(t, A) = \mu(X^{-1}(t, A)) = 0$. D'où le résultat pour $t \in [0, 1[$. Lorsque $t = 1$, $\bar{\rho}_p(t, \cdot) = \nu$ qui est absolument continue. En utilisant la proposition 2.1 pour tout $t \in [0, 1]$ avec les mesures μ et $\bar{\rho}_p(t, \cdot)$, on a l'existence de $X(t, \cdot)^{-1}$ presque partout et on peut poser

$$v_p(t, \cdot) = (\partial_t X)(t, X^{-1}(t, \cdot)) = \nabla\psi(X^{-1}(t, \cdot)) - X^{-1}(t, \cdot).$$

Ainsi

$$\begin{aligned} \int_0^1 \int_D |v_p|^2 d\bar{\rho}_p(t, x) dt &= \int_0^1 \int_D |v_p(t, X(t, x))|^2 d\mu(x) dt \\ &= \int_0^1 \int_D |\nabla\psi(X^{-1}(t, X(t, x))) - X^{-1}(t, X(t, x))|^2 d\mu(x) dt \\ &= \int_0^1 \int_D |\nabla\psi(x) - x|^2 d\mu(x) dt = W_2(\mu, \nu)^2. \end{aligned}$$

Il reste encore à vérifier que $(\bar{\rho}_p, v_p) \in V(\mu, \nu)$. On a

- $t \mapsto \bar{\rho}_p(t, \cdot)$ est dans $C^0([0, 1], \mathcal{P}^{\text{ac}}(D))$, car $\bar{\rho}_p$ est continue en t par construction et $\bar{\rho}_p(t, \cdot)$ est a.c.
- v_p est dans $L^2(\bar{\rho}_p(t, \cdot) dt)$, car $\int_0^1 \int_D |v_p|^2 d\bar{\rho}_p(t, x) dt = W_2(\mu, \nu)^2 < +\infty$ car D est borné,
- $(\bar{\rho}_p, v_p)$ est une solution de l'équation de la continuité au sens faible, car théorème 2.1 et $\bar{\rho}_p(0, \cdot) = \mu$ et $\bar{\rho}_p(1, \cdot) = \nu$ dans D , par construction.

• Conclusion :

On devait faire l'étape 2 car on a pas de régularité sur v_p dans l'étape 3. On a ainsi montré que

$$W_2(\mu, \nu)^2 \leq \inf_{(\bar{\rho}, v) \in V(\mu, \nu)} \int_0^1 \int_D |v(t, x)|^2 d\bar{\rho}(t, x) dt \leq \int_0^1 \int_D |v_p(t, x)|^2 d\bar{\rho}_p(t, x) dt = W_2(\mu, \nu)^2.$$

D'où le résultat. □

3 La méthode du Lagrangien (Augmenté)

Cette section rappelle des résultats donnés dans [2]. On considère le problème d'optimisation suivant :

$$\min_{v \in V} F(Bv) + G(v), \tag{2}$$

avec

- V et H deux espaces de Hilbert,
- B une application linéaire de V dans H ,
- F et G des fonctions convexes, propres et s.c.i sur H et V respectivement.

On pose

$$J(v) := F(Bv) + G(v).$$

Théorème 3.1. *Le problème (2) admet une solution si J est coercive i.e.*

$$\lim_{\|v\|_V \rightarrow +\infty} J(v) = +\infty.$$

De plus, si J est strictement convexe et propre, alors on a l'unicité.

3.1 Le Lagrangien (Augmenté)

Le problème (2) est équivalent au problème suivant :

$$\min_{(v, q) \in W} F(q) + G(v), \tag{3}$$

avec

$$W = \{(v, q) \in V \times H \mid Bv - q = 0\}.$$

On peut définir le *Lagrangien*, $\mathcal{L} : V \times H \times H \rightarrow \bar{\mathbb{R}}$ par,

$$\mathcal{L}(v, q, \mu) = F(q) + G(v) + \langle \mu, Bv - q \rangle_H.$$

Définition 3.1. *On dit que (u, p, λ) est un point-selle de \mathcal{L} sur $V \times H \times H$ si*

$$\mathcal{L}(u, p, \mu) \leq \mathcal{L}(u, p, \lambda) \leq \mathcal{L}(v, q, \lambda), \quad \forall (v, q) \in V \times H, \quad \forall \mu \in H.$$

Proposition 3.1. *(u, p, λ) est un point-selle de \mathcal{L} sur $V \times H \times H$, si et seulement si,*

$$\inf_{(v, q) \in V \times H} \sup_{\mu \in H} \mathcal{L}(v, q, \mu) = \sup_{\mu \in H} \inf_{(v, q) \in V \times H} \mathcal{L}(v, q, \mu) = \mathcal{L}(u, p, \lambda).$$

On a le résultat suivant,

Théorème 3.2. Si (u, p, λ) est un point-selle de \mathcal{L} sur $V \times H \times H$, alors u est une solution de (2) et $p = Bu$.

Remarque 8. En dimension infinie, l'existence d'un point-selle (et donc d'un multiplicateur de Lagrange λ) est problématique. En effet, il existe des conditions pour l'existence d'un point-selle mais celles-ci sont rarement satisfaites en pratique. En dimension finie, on a l'existence d'un point-selle si le problème (2) admet une solution.

En particulier, pour les problèmes qui sont discrétisés à partir de problèmes de dimension infinie, on peut avoir u_h qui tends vers u lorsque $h \rightarrow 0$ mais λ_h ne converge pas dans H .

On peut également définir le *Lagrangien augmenté* $\mathcal{L}_r : V \times H \times H$ par,

$$\mathcal{L}_r(v, q, \mu) = F(q) + G(v) + \langle \mu, Bv - q \rangle_H + \frac{r}{2} \|Bv - q\|_H^2.$$

Proposition 3.2. Tout points-selle de \mathcal{L} est un point-selle de \mathcal{L}_r , $\forall r > 0$, et vice-versa.

Remarque 9. L'ajout du terme $\frac{r}{2} \|Bv - q\|_H^2$ permet uniquement d'améliorer la convergence des algorithmes.

3.2 Un algorithme de Uzawa

On considère l'algorithme suivant :

Pour $(p^{n-1}, \lambda^n) \in H \times H$ donné,

— Étape 1 : on pose

$$u^n := \operatorname{argmin}_{v \in V} \mathcal{L}_r(v, q^{n-1}, \lambda^n).$$

— Étape 2 : on pose

$$p^n := \operatorname{argmin}_{q \in H} \mathcal{L}_r(u^n, q, \lambda^n).$$

— Étape 3 : on pose

$$\lambda^{n+1} := \lambda^n + \rho_n (Bu^n - p^n).$$

— Retour à l'étape 1.

Théorème 3.3. On suppose que \mathcal{L}_r admet un point-selle (u, p, λ) sur $V \times H \times H$. Si $0 < \rho_n < \frac{1+\sqrt{5}}{2}r$, alors

- $u^n \rightarrow u$ dans V ,
- $p^n \rightarrow p$ dans H ,
- $\lambda^{n-1} - \lambda^n \rightarrow 0$ dans H ,
- λ^n est bornée dans H .

Remarque 10. Puisque

$$\lambda^{n+1} := \lambda^n + \rho_n (Bu^n - p^n),$$

le résultat $\lambda^{n-1} - \lambda^n \rightarrow 0$ devient $(Bu^n - p^n) \rightarrow 0$ dans H .

4 Le Lagrangien augmenté pour la formule de Benamou-Brenier

On prend $V = \{(\rho, m) : [0, 1] \times D \rightarrow \mathbb{R} \times \mathbb{R}^d \mid \exists(\bar{\rho}, v) \in V(\mu, \nu), \bar{\rho} \text{ a.c. à densité } \rho > 0 \text{ et } m = \rho v\}$ et $H = L^2([0, 1] \times D, \mathbb{R})$, et on pose

$$\langle u, v \rangle_V = \int_0^1 \int_D u(t, x) \cdot v(t, x) \, dx \, dt,$$

$$\langle u, v \rangle_H = \int_0^1 \int_D u(t, x) v(t, x) \, dx \, dt.$$

Soit $(\rho, m) \in V$, on pose $v := \frac{m}{\rho}$. Alors,

$$\int_0^1 \int_D \rho(t, x) |v(t, x)|^2 \, dx \, dt = 2 \int_0^1 \int_D \frac{|m(t, x)|^2}{2\rho(t, x)} \, dx \, dt.$$

On peut alors réécrire le problème d'optimisation de la formule de Benamou-Brenier sous la forme du problème d'optimisation de la Section 3 :

$$\inf_{(\rho, m) \in V} G(\rho, m) + F(B(\rho, m)), \quad (4)$$

avec

$$G(\rho, m) := \int_0^1 \int_D \frac{|m(t, x)|^2}{2\rho(t, x)} dx dt,$$

$$F(q) := \begin{cases} 0, & \text{si } q = 0, \\ +\infty, & \text{sinon,} \end{cases}$$

$$B(\rho, m) := \partial_t \rho + \operatorname{div} m.$$

En utilisant les résultats de la section précédente, on a que $(\rho, m) \in V$ solution de (4) si et seulement si, $((\rho, m), q, \phi) \in V \times H \times H$ point-selle de

$$F(q) + G(\rho, m) + \langle \phi, B(\rho, m) - q \rangle_H. \quad (5)$$

Or lorsque q n'est pas identiquement nulle, F vaut $+\infty$. On doit donc avoir que $q = 0$. Ainsi $((\rho, m), \phi) \in V \times H \times H$ point-selle de (5) si et seulement si, $((\rho, m), \phi) \in V \times H$ point-selle de

$$G(\rho, m) + \langle \phi, B(\rho, m) \rangle_H.$$

Cela signifie que le problème (4) est équivalent au problème

$$\inf_{(\rho, m) \in V} \sup_{\phi \in H} G(\rho, m) + \langle \phi, B(\rho, m) \rangle_H. \quad (6)$$

Le problème de cette formulation est le manque de convexité de G par rapport à (ρ, m) , ce qui ne nous permet pas d'utiliser les résultats du Lagrangien de la section précédente pour la résolution de ce problème de minimisation. Dans la suite, on va transformer le problème (6) afin d'obtenir un nouveau problème ayant cette fois-ci les bonnes propriétés. On commence par donner le lemme suivant :

Lemme 4.1.

$$\frac{|m(t, x)|^2}{2\rho(t, x)} = \sup_{(a, b) \in K} a(t, x)\rho(t, x) + b(t, x) \cdot m(t, x)$$

avec

$$K := \left\{ (a, b) : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R} \times \mathbb{R}^d \mid a(t, x) + \frac{|b(t, x)|^2}{2} \leq 0, \forall (t, x) \in \mathbb{R} \times \mathbb{R}^d \right\}.$$

Démonstration. Soit b dans \mathbb{R}^d . Pour que (a, b) soit dans K , il faut que $a \leq -\frac{|b|^2}{2}$. Puisque $\rho > 0$, on peut chercher d'abord à maximiser a . On prend $a = -\frac{|b|^2}{2}$. Ainsi,

$$\sup_{(a, b) \in K} a(t, x)\rho(t, x) + b(t, x) \cdot m(t, x) = \sup_{b \in \mathbb{R}^d} \underbrace{-\frac{|b(t, x)|^2}{2}\rho(t, x) + b(t, x) \cdot m(t, x)}_{\text{strictement concave}}.$$

On calcul le gradient :

$$-b\rho + m = 0 \Leftrightarrow b = \frac{m}{\rho}.$$

Ainsi,

$$\sup_{(a, b) \in K} a(t, x)\rho(t, x) + b(t, x) \cdot m(t, x) = \sup_{b \in \mathbb{R}^d} -\frac{|\frac{m}{\rho}(t, x)|^2}{2}\rho(t, x) + \frac{m}{\rho}(t, x) \cdot m(t, x) = \frac{|m(t, x)|^2}{2\rho(t, x)}.$$

□

On a donc que

$$G(\rho, m) + \langle \phi, B(\rho, m) \rangle_H = \int_0^1 \int_D \sup_{(a,b) \in K} a(t,x)\rho(t,x) + b(t,x) \cdot m(t,x) dx dt - \int_D \int_0^1 (\rho \partial_t \phi + m \nabla_x \phi) dx dt - \underbrace{\int_D (\phi(0,x)\mu(x) - \phi(1,x)\nu(x)) dx}_{=: \tilde{G}(\phi)}. \quad (7)$$

On pose $\gamma = (\rho, m)$ et $q := (a, b)$ et on a,

$$\begin{aligned} (7) &= \sup_{q \in K} \langle q, \gamma \rangle_V - \langle \gamma, \nabla_{t,x} \phi \rangle_V - \tilde{G}(\phi) = \sup_q -\tilde{F}(q) - \langle \gamma, \nabla_{t,x} \phi - q \rangle_V - \tilde{G}(\phi) \\ &= -\inf_q \tilde{F}(q) + \langle \gamma, \nabla_{t,x} \phi - q \rangle_V + \tilde{G}(\phi), \end{aligned}$$

avec

$$\tilde{F}(q) := \begin{cases} 0, & \text{si } q \in K, \\ +\infty, & \text{sinon.} \end{cases}$$

Ainsi,

$$(6) \Leftrightarrow \sup_{\gamma} \inf_{(\phi, q)} \tilde{F}(q) + \tilde{G}(\phi) + \langle \gamma, \nabla_{t,x} \phi - q \rangle_V.$$

Donc ϕ joue le rôle de l'inconnue et γ joue le rôle du Lagrangien pour la contrainte $\nabla_{t,x} \phi = q$. On peut définir le Lagrangien augmenté et obtenir

$$L_r(\phi, q, \gamma) := \tilde{G}(\phi) + \tilde{F}(q) + \langle \gamma, \nabla_{t,x} \phi - q \rangle_V + \frac{r}{2} \|\nabla_{t,x} \phi - q\|^2.$$

5 Algorithmes d'Uzawa

Maintenant on utilise les résultats de la Section 3 sur le Lagrangien L_r .

5.1 Étape 1

On commence par la première étape de l'algorithme d'Uzawa donné dans la section sur le Lagrangien :

$$\begin{aligned} \phi^n &:= \operatorname{argmin}_{\phi} L_r(\phi, q^{n-1}, \gamma^n) \\ &= \operatorname{argmin}_{\phi} \tilde{F}(q^{n-1}) + \tilde{G}(\phi) + \langle \gamma^n, \nabla_{t,x} \phi - q^{n-1} \rangle + \frac{r}{2} |\nabla_{t,x} \phi - q^{n-1}|^2 \\ &= \operatorname{argmin}_{\phi} \tilde{G}(\phi) + \langle \gamma^n, \nabla_{t,x} \phi - q^{n-1} \rangle + \frac{r}{2} |\nabla_{t,x} \phi - q^{n-1}|^2. \end{aligned}$$

Puisque L_r est différentiable par rapport à ϕ , on peut calculer l'équation d'Euler-Lagrange (long mais pas très difficile) et obtenir l'équation suivante : for ϕ in H ,

$$\tilde{G}(\phi) + r \langle \nabla_{t,x} \phi^n - q^{n-1}, \nabla_{t,x} \phi \rangle_V + \langle \gamma^n, \nabla_{t,x} \phi \rangle_V = 0.$$

Cette équation est la formulation faible de l'EDP suivante :

$$\begin{cases} -r \Delta_{t,x} \phi = \nabla_{t,x} \cdot (m^n - q^{n-1}), & \text{dans } [0, 1] \times D, \\ r \partial_t \phi^n(0, \cdot) = \mu - \rho^n(0, \cdot) + r a^{n-1}(0, \cdot), & \text{dans } D, \\ r \partial_t \phi^n(1, \cdot) = \nu - \rho^n(1, \cdot) + r a^{n-1}(1, \cdot), & \text{dans } D, \\ \text{conditions périodiques,} & \text{sur } [0, 1] \times \partial D. \end{cases}$$

5.2 Étape 2

Contrairement à l'étape 1, on ne peut pas utiliser l'équation d'Euler-Lagrange pour résoudre le problème d'optimisation car \tilde{F} n'est pas différentiable en q . Cependant,

$$\begin{aligned}
 q^n &:= \operatorname{argmin}_{q \in K} L_r(\phi^n, q, \gamma^n) \\
 &= \operatorname{argmin}_{q \in K} \underbrace{\tilde{F}(q)}_{=0} + \underbrace{\tilde{G}(\phi^n)}_{=0} + \langle \gamma^n, \nabla_{t,x} \phi^n - q \rangle_H + \frac{r}{2} |\nabla_{t,x} \phi^n - q|^2 \\
 &= \operatorname{argmin}_{q \in K} 2 \underbrace{\langle \frac{\gamma^n}{r}, \nabla_{t,x} \phi^n - q \rangle_H}_{=0} + |\nabla_{t,x} \phi^n - q|^2 + \underbrace{\left| \frac{\gamma^n}{r} \right|^2}_{=0} \\
 &= \operatorname{argmin}_{q \in K} \left| \nabla_{t,x} \phi^n - q + \frac{\gamma^n}{r} \right|^2.
 \end{aligned}$$

Alors, $q(t, x) = (a(t, x), b(t, x))$ est la projection de $\nabla_{t,x} \phi^n(t, x) + \frac{\gamma^n(t, x)}{r}$ sur K . Dû à la forme particulière de K , il s'agit en fait d'une projection unidimensionnelle et il est possible de l'exprimer de manière analytique en utilisant les formules de Cardan (un fichier d'illustration GeoGebra est disponible). Il est important de noter que cette projection est ponctuelle.

5.3 Étape 3

Enfin la dernière étape est une simple mise-à-jour ponctuelle de la variable μ :

$$\gamma^{n+1}(t, x) := \gamma^n(t, x) + r (\nabla_{t,x} \phi^n(t, x) - q^n(t, x)).$$

5.4 Condition d'arrêt

Dans le théorème de la convergence de l'algorithme, on avait en particulier que $(B\mu^n - p^n) \rightarrow 0$ dans H . Dans notre cas, cela correspond à $B\phi^n - p^n = \nabla_{t,x} \phi^n(t, x) - q^n$

6 Résultats numériques

On utilise les différences finies pour la résolution de l'EDP pour l'étape 1 de l'algorithme. Voici quelques résultats numériques (un notebook jupyter est disponible) :

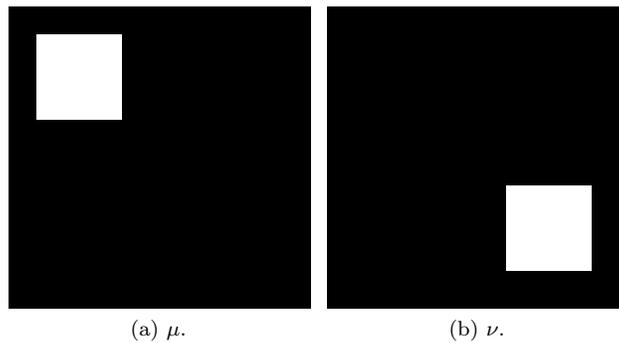


FIGURE 3 – Densité initiale et finale.

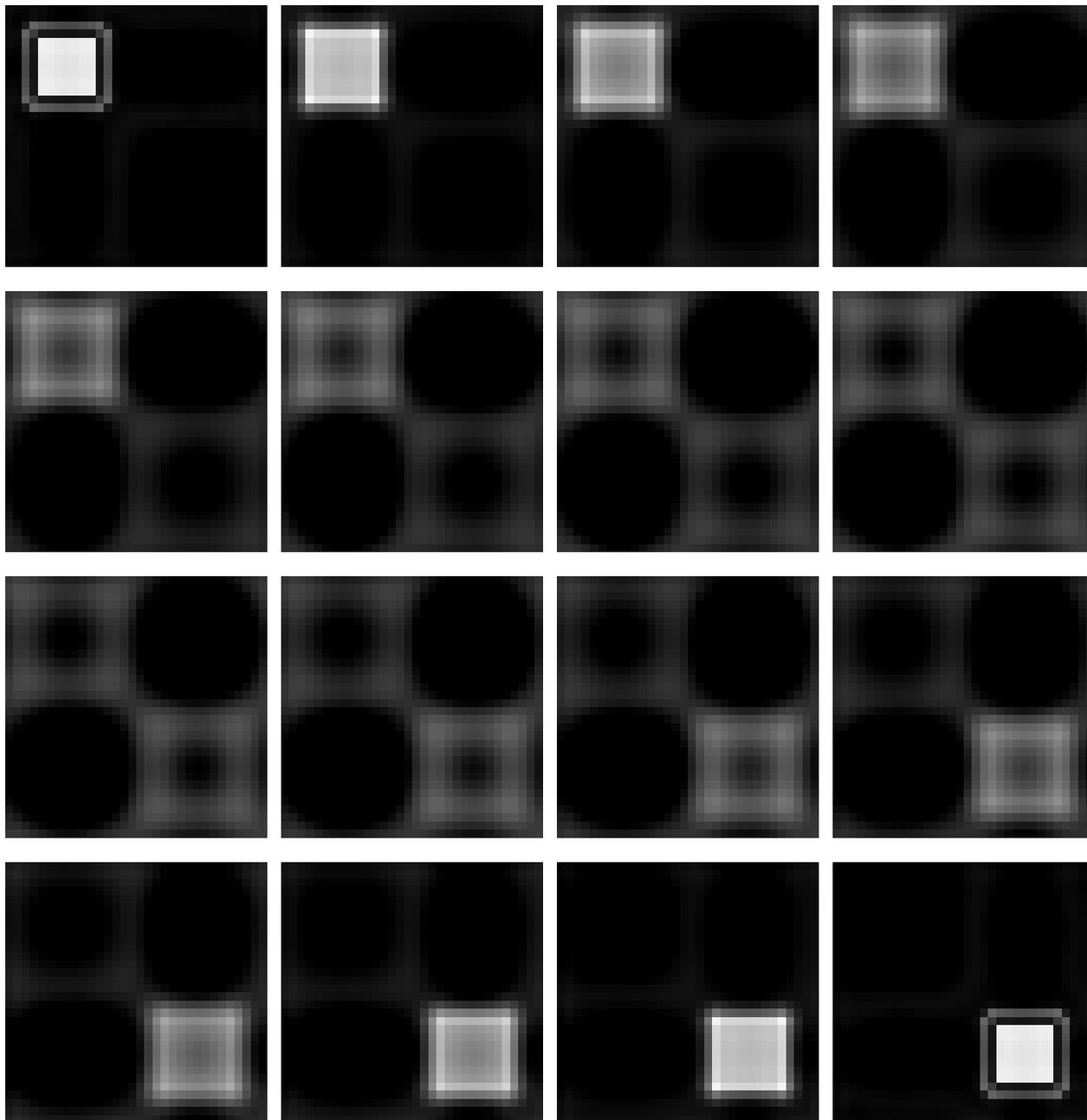


FIGURE 4 – Densités $\rho(t, \cdot)$.

Références

- [1] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. 84(3) :375–393.
- [2] Michel Fortin and Roland Glowinski. *Augmented Lagrangian Methods : Applications to the Numerical Solution of Boundary-Value Problems*. Elsevier Science. OCLC : 1049710365.
- [3] Filippo Santambrogio. *Optimal transport for applied mathematicians : calculus of variations, PDEs, and modeling*. Number volume 87 in Progress in nonlinear differential equations and their applications. Birkhäuser.
- [4] Cédric Villani. *Topics in optimal transportation*. Number v. 58 in Graduate studies in mathematics. American Mathematical Society.