DRAFT : CONVEXITE CACHEE DANS LES EDP

Yann Brenier...

March 3, 2020

1 Quelques exemples simples de convexité cachée en dehors des EDP

1.1 Deux premiers exemples très élémentaires

Theorem 1. *Soit K un espace métrique compact et f une fonction continue sur K On note $P(K)$ l'ensemble des mesures de probabilité boréliennes sur K. Alors, il est équivalent de dire que f atteint son minimum en x_0 dans K et que δ_{x_0} réalise sur $P(K)$ le minimum de la fonctionnelle linéaire*

$$\mu \in P(K) \to F(\mu) = \int_K f(x) d\mu(x)$$

Preuve. Si x_0 réalise le minimum de f sur K, alors pour toute $\mu \in P(K)$ on a

$$F(\mu) \geq \int_K f(x_0) d\mu(x) = f(x_0)$$

et d'autre part

$$F(\delta_{x_0}) = f(x_0).$$

Donc δ_{x_0} réalise bien le minimum de F sur $P(K)$. Réciproquement, si δ_{x_0} réalise le minimum de F sur $P(K)$, on a pour tout $x \in K$,

$$f(x_0) = F(\delta_{x_0}) \leq F(\delta_x) = f(x),$$

ce qui montre bien que le minimum de f est atteint en x_0.
Remarque : notons que si le minimum de f est atteint en plusieurs points distincts $x_0, \cdots, x_N$ alors le minimum de F est atteint par n'importe quelle combinaison convexe des δ_{x_i}.
Remarque : le résultat reste valable même si f est seulement l.s.c de K dans $]-\infty, +\infty]$ non identiquement égale à $+\infty$. Notons qu'alors la fonctionnelle F ne peut plus être considérée comme linéaire mais seulement comme une fonctionnelle l.s.c. (relativement à la convergence faible-* de $P(K)$) à valeurs dans $]-\infty, +\infty]$ non identiquement égale à $+\infty$.

Theorem 2. *Soit H un espace de Hilbert séparable de dimension infinie. Alors, la boule unité de H est la fermeture faible de la sphère unité.*

Remarque : le résultat est évidemment faux en dimension finie.
Preuve : Comme on est en dimension infinie on peut construire une suite infinie de vecteurs $u_n \in H$ orthonormés, i.e. tels que $(u_n|u_m) = \delta_{nm}$. Cette suite tend faiblement vers 0. En effet, pour tout $x \in H$, on a :

$$0 \leq |x - \sum_{i=1}^{N}(x|u_n)u_n|^2 = |x|^2 - \sum_{i=1}^{N}(x|u_n)^2.$$

Donc la série des $(x|u_n)$ est de carré sommable. Ainsi, son terme générique $(x|u_n)$ tend vers zéro, ce qui prouve bien que u_n tend vers 0 faiblement. Soit maintenant x tel que $|x| \leq 1$. Pour tout n, posons $x_n = x + r_n u_n$ où $r_n \in \mathbb{R}$ est choisi de sorte que $|x_n| = 1$. Ceci est possible, car il suffit que

$$|x|^2 + 2r_n(x|u_n) + r_n^2 = 1,$$

i.e.

$$(r_n + (x|u_n))^2 = 1 - |x|^2 + (x|u_n)^2,$$

et donc

$$r_n = -(x|u_n) + \sqrt{1 - |x|^2 + (x|u_n)^2}$$

convient (puisque $|x| \leq 1$). Notons que, du coup,

$$|r_n| \leq |x| + 1,$$

ce qui montre que, quitte à extraire une sous-suite, on peut supposer que $r_n \to r$, pour un réel r. Ainsi on a trouvé un point x_n de la sphère unité qui converge faiblement vers x. En effet, pour tout $y \in H$, on a

$$(x_n - x|y) = (r_n u_n|y) = (r_n - r)(u_n|y) + r(u_n|y)$$

où $|(r_n - r)(u_n|y)| \leq |r_n - r||y| \to 0$ et $(u_n|y) \to 0$ puisque u_n converge faiblement vers 0. Ainsi, on peut approcher faiblement tout point de la boule unité par une suite faiblement convergente de la sphère unité...à condition d'avoir "suffisamment de place" grace à la dimension infinie !

1.2 Un exemple un peu moins élémentaire

Theorem 3. *Soit DS_N l'ensemble des matrices $N \times N$ réelles à valeurs positives ou nulles, dont la somme de chaque colonne et de chaque ligne vaut* 1. *(On dit d'une telle matrice qu'elle est bistochastique). Alors DS_N est exactememt l'enveloppe convexe des matrices de permutation, i.e. de celles des matrices bistochastiques à valeurs dans $\{0, 1\}$.*

Que l'enveloppe convexe des matrices de permutation soit incluse dans DS_N est immédiat. La réciproque, due à G. Birkhoff, est une conséquence assez directe du fameux "lemme du mariage" en combinatoire.
Il a une conséquence importante en optimisation combinatoire :

Theorem 4. *Soit c_{ij} une matrice réelle $N \times N$ fixée. Alors il est équivalent de résoudre :*
1) Le problème d'affectation "linéaire" :

$$\inf_{\sigma \in \mathfrak{S}_N} \sum_{i=1,N} c_{i\sigma_i}$$

où $\mathfrak{S}_N$ désigne le groupe symétrique (i.e. le groupe des permutations des N premiers entiers) ;
2) Le "programme linéaire"

$$\inf_{s \in DS_N} \sum_{i,j=1}^{N} c_{ij} s_{ij}.$$

Cet exemple est frappant car il réduit un problème d'optimisation combinatoire à un simple programme linéaire (i.e. à la minimisation d'une fonctionnelle linéaire avec contraintes linéaires d'égalité ou d'inégalité).
Remarque : il existe des algorithmes de coùt $O(N^3)$ pour ce problème [15] considéré comme trés simple en optimisation combinatoire. En revanche, il y a des problèmes d'optimisation combinatoire connus pour ne pas pouvoir être réduits à un problème d'optimisation convexe, comme le problème d'affectation "quadratique" pour lequel une seconde matrice carrée γ_{ij} est donnée et où il s'agit de résoudre :

$$\inf_{\sigma \in \mathfrak{S}_N} \sum_{i,j=1,N} c_{ij} \gamma_{\sigma_i \sigma_j}.$$

(Ce problème est NP. Un cas particulier n'est rien d'autre que le fameux problème du voyageur de commerce.) Néanmoins, dans certains cas, ce problème peut se ramener à la recherche de points d'équilibre de systèmes dynamiques admettant une formulation convexe [65].

1.3 Une version "continue" du théorème de G. Birkhoff

Considérons le cube unité $D = [0,1]^d$. On peut le découper en $N = 2^{nd}$ sous-cubes de volumes égaux D_α pour $\alpha = 1, \cdots, N$ et attacher à chaque permutation $\pi \in \mathfrak{S}_N$ la transformation T_π du cube dans lui-même qui translate en bloc chaque cube D_α vers $D_{\pi(\alpha)}$. Ainsi on a défini, au moins Lebesgue presque partout, une application qui conserve la mesure de Lebesgue $\mathcal{L}$, restreinte à D. Cela fait de T_π un élément de l'ensemble, noté $VPM(D)$, de toutes les applications boréliennes T de D dans lui-même qui conserve le volume, i.e. la mesure de Lebesgue $\mathcal{L}_D$, au sens que :

$$\mathcal{L}(T^{-1}(A)) = \mathcal{L}(A),$$

pour tout borélien de D, ce qu'on écrit en bref $\mathcal{L} \circ T^{-1} = \mathcal{L}$, ou, de façon équivalente,

$$\int_D f(T(x))dx = \int_D f(x)dx,$$

pour toute fonction $f \in C(\mathbb{R}^d)$, i.e. pour toute $f : \mathbb{R}^d \to \mathbb{R}$ continue. L'étude détaillée de $VPM(D)$ sera développée dans l'appendice 15 pour le pas perdre le fil de l'exposé. D'ores et déjà, on peut énoncer quelques propriétés de $TCM(D)$, toutes démontrées dans l'appendice 15
1) L'ensemble $VPM(D)$ peut être vu comme une partie fermée de l'espace de Hilbert $H = L^2(D; \mathbb{R}^d)$, contenue dans une sphère et qui n'est ni convexe ni compacte. C'est aussi un semi-groupe pour la loi de composition des applications, mais pas un groupe, puisqu'en général, les applications de $VPM(D)$ ne sont pas inversibles, comme le montre l'exemple simple, quand $d = 1$, de l'application

$$T(x) = 2x \text{ mod.}{-1}\ .$$

2) $VPM(D)$ contient comme sous-ensemble le groupe des transformations qui sont bijectives à un ensemble Lebesgue négligeable près, et donc inversibles, mais ce groupe a le mauvais goût de ne pas être fermé dans H et sa complétion est en fait $VPM(D)$ lui-même.
3) $VPM(D)$ contient aussi le sous-groupe $P_N(D)$ des "permutations", i.e. des application T_π obtenues comme on l'a vu plus haut par décomposition dyadique de D en $N = 2^{nd}$ sous-cubes, pour toutes les permutations $\pi \in \mathfrak{S}_N$. L'union des $P_N(D)$ pour toutes les valeurs entières de $n \in \mathbb{N}$ forme encore un sous-groupe de $VPM(D)$, noté $P(D)$, dont on peut montrer qu'il est dense dans $VPM(D)$.
4) $VPM(D)$ contient le groupe, souvent noté $SDiff(D)$, de tous les difféomorphismes T conservant l'orientation et le volume de D, au sens que chaque T est la restriction à D d'un difféomorphisme de $\mathbb{R}^d$ telle que $T(D) = D$ et

$$\det(DT(x)) = 1, \quad \forall x \in \ D.$$

Ce groupe est trivial quand $d = 1$ (il est alors réduit à l'application identité) et sa complétion est $VPM(D)$ tout entier dès que $d \geq 2$.
Ce dernier résultat a la conséquence troublante qu'un difféomorphisme de D qui conserve le volume mais renverse l'orientation, par exemple, quand $d = 2$,

$$T(x_1, x_2) = (1 - x_1, x_2), \ \ \forall x = (x_1, x_2) \in [0, 1]^2,$$

peut être approché (au sens L^2 "fort") par une suite de difféomorphismes conservant le volume *et* l'orientation!
Comme il a été mentionné, $VPM(D)$ est certes une partie fermée bornée du Hilbert $H = L^2(D; \mathbb{R}^d)$, qui n'est ni convexe ni compacte. Cela dit, on a l'équivalent du théorème de Birkhoff, au sens qu'on peut compactifier $VPM(D)$ en un ensemble convexe, l'ensemble des mesures bistochastiques de D, noté $DS(D)$ et défini comme l'ensemble des mesures de probabilité boréliennes μ sur le produit $D \times D$ dont chaque projection est la mesure de Lebesgue sur D, ce qui veut dire

$$\mu(D \times A) = \mu(A \times D) = \mathcal{L}(A),$$

pour tout borélien A de D, ou, de façon équivalente, (i.e. pour toute fonction continue f sur D). L'injection

$$T \in VPM(D) \to \mu_T \in DS(D),$$

se fait simplement en posant

$$\int_{D\times D} f(x,y)d\mu(x,y) = \int_D f(x,T(x))dx, \quad \forall f \in C^0(D\times D).$$

et on montre (voir Appendice 15):

Theorem 5. *Dès que $d \geq 2$, l'espace $DS(D)$ des mesures bistochastiques μ sur $D = [0,1]^d$, i.e. des mesures de Borel positives sur D^2 telles que:*

$$\int_{D\times D} f(x)d\mu(x,y) = \int_{D\times D} f(y)d\mu(x,y) = \int_D f(x)dx, \quad \forall f \in C^0(D)$$

est exactement la complétion du groupe $SDiff(D)$ des difféomorphismes de D qui conservent le volume et l'orientation, au sens de la convergence faible des mesures. i.e. au sens de la topologie faible- sur le dual de l'espace de Banach (pour la norme du sup) des fonctions continues sur le produit $D\times D$, i.e. $C^0(D\times D)$. Autrement dit, tout $\mu \in DS(D)$ admet une suite (T_n) d'éléments de $SDiff(D)$ telles que*

$$\int_{D\times D} f(x,y)d\mu(x,y) = \lim_n \int_D f(x,T_n(x))dx, \quad \forall f \in C^0(D\times D).$$

Ceci peut être vu comme l'équivalent "continu" du théorème de Birkhoff ou la notion de convergence faible se substitue à celle d'enveloppe convexe.
Ainsi on a une nouvelle illustration, cette fois dans le cadre continu, du concept de convexité cachée : dès que $d \geq 2$, se cache dans le groupe $SDiff(D)$ des difféomorphismes du cube unité $D = [0,1]^d$ conservant le volume et l'orientation, la structure convexe de l'ensemble $DS(D)$ des mesures bistochastiques sur $D\times D$, tout comme derrière le groupe des permutations $\mathfrak{S}_N$ se cachait l'ensemble convexe des matrices bistochastiques $N\times N$.
Notons aussi la similarité avec le théorème 2, $DS(D)$ jouant le rôle de la boule unité et $SDiff(D)$ celui de la sphère unité.

1.4 Permutations et convexité liées par la géométrie euclidienne

Avec le théorème de G. Birkhoff, nous avons déja relié les concepts de convexité et de permutation. Voici un autre lien induit par la géométrie euclidienne. Considérons deux ensembles de N points distincts dans l'espace euclidien $\mathbb{R}^d$, respectivement notés $A(1),\cdots,A(N)$ et $B(1),\cdots,B(N)$. On appelle nuages de points $\mathcal{A}$, $\mathcal{B}$ ces deux ensembles qu'on peut fructueusement considérer comme des listes de points modulo leur étiquetage, i.e. $\mathcal{A} \in (\mathbb{R}^d)^N/\mathfrak{S}_N$, $\mathcal{B} \in (\mathbb{R}^d)^N/\mathfrak{S}_N$. Il est alors naturel de définir le carré de leur distance par

$$d(\mathcal{A},\mathcal{B})^2 = \inf_{\sigma\in\mathfrak{S}_n} \frac{1}{N} \sum_{i=1,N} |A(\sigma_i) - B(i)|^2,$$

où $|\cdot|$ est la norme euclidienne sur $\mathbb{R}^d$. On a alors

Theorem 6. *Dans la définition de la distance entre les deux nuages de points $\mathcal{A}$ et $\mathcal{B}$, la permutation identité est optimale si et seulement s'il existe une fonction convexe $\Phi : \mathbb{R}^d \to \mathbb{R}$, qu'on peut supposer C^1, telle que $B(i) = (\nabla\Phi)(A(i))$, $\forall i \in \{1, \cdots, N\}$.*

Corrélations maximales entre variables aléatoires de lois données

Ce théorème admet une version continue (et même de dimension infinie!), due à Knott et Smith [164], en termes de lois de probabilité et de "variables aléatoires" :

Theorem 7. *Soient $(H, ((\cdot,\cdot)))$ un espace de Hilbert séparable et deux mesures de probabilités de Borel μ et ν sur H de second moment fini, i.e.*

$$\int_H ((x,x))^2 d\mu(x) < +\infty, \int_H ((x,x))^2 d\nu(x) < +\infty.$$

Soient X et Y deux variables aléatoires de lois respectives μ et ν sur H, alors la corrélation entre elles, i.e. $((X,Y))$ est maximale si et seulement s'il existe une fonction convexe $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ l.s.c sur H telle que, presque sûrement,

$$(X, Y) \in \partial\Phi$$

où $\partial\Phi$ est le sous-différentiel de Φ défini par :

$$\partial\Phi = \{(x,\xi) \in H \times H,\ \Phi(y) \geq \Phi(x) + ((\xi, y-x)),\ \forall y \in H\}.$$

Remarque : Au lieu de corrélation maximale, on aurait pu parler de distance au carré $((X-Y, X-Y))$ minimale, puisque les termes carrés $((X,X))$ et $((Y,Y))$ sont respectivement déterminés par les lois μ et ν.
Le théorème précédent n'est ainsi que la version discrète de ce dernier, en termes de "nuages de points" plutôt que de variables aléatoires de loi fixée. Le caractère discret nous autorise dès lors à supposer Φ lisse plutôt que seulement l.s.c.

Transport optimal de mesures

Un résultat intermédiaire dit "de transport optimal de mesures", cette fois en dimension finie mais plus propre à l'analyse des EDP, s'écrit

Theorem 8. *Soient deux mesures de probabilités de Borel μ et ν sur $\mathbb{R}^d$ de second moment fini, i.e.*

$$\int_{\mathbb{R}^d} |x|^2 d\mu(x) < +\infty, \int_{\mathbb{R}^d} |x|^2 d\nu(x) < +\infty.$$

Si on suppose μ absolument continue par rapport à la mesure de Lebesgue, alors il existe une application borélienne $T : \mathbb{R}^d \to \mathbb{R}^d$, transportant μ vers ν au sens que $\nu = \mu \circ T^{-1}$, ou encore

$$\int_{\mathbb{R}^d} f(T(x)) d\mu(x) = \int_{\mathbb{R}^d} f(y) d\nu(y), \quad \forall f \in C_c^0(\mathbb{R}^d),$$

qui admet un "potentiel convexe" au sens qu'il existe une fonction convexe l.s.c $\Phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ telle que, pour presque tout point x, Φ est différentiable en x avec $T(x) = \nabla\Phi(x)$.
Cette application T est uniquement définie $\mu-$presque partout sur $\mathbb{R}^d$ et, de plus, est caractérisée comme l'unique minimiseur de la fonctionnelle

$$\int_{\mathbb{R}^d} |T(x) - x|^2 d\mu(x)$$

parmi toutes les applications boréliennes transport μ vers ν.

Remarque : La condition d'absolument continuité est certainement restictive (elle élimine la plupart des mesures μ ayant des atomes, par exemple), mais apporte en revanche un précieux résultat d'unicité.

Une décomposition de Helmholtz non-linéaire : la factorisation polaire des applications

Dans le cas particulier μ est la mesure de Lebesgue sur un domaine compact D de $\mathbb{R}^d$ (normalisée de sorte que la mesure de D soit 1), on peut rephraser le résultat précédent en termes de "factorisation polaire d'applications", dans un langage plus proche du cas discret vu plus haut :

Theorem 9. *Soit D un domaine compact de $\mathbb{R}^d$ et $\mathcal{L}$ la mesure de Lebesgue sur D. Soit $Y \in L^2(D, \mathbb{R}^d)$ telle que $\mathcal{L} \circ Y^{-1} << \mathcal{L}$, alors il existe une unique "factorisation polaire" de Y sous la forme $Y = T \circ X$ où*
1) $X : D \to D$ conserve la mesure de Lebesgue : $\mathcal{L} \circ X^{-1} = \mathcal{L}$;
2) $T : D \to \mathbb{R}^d$ admet un potentiel convexe, au sens qu'il existe une fonction convexe l.s.c $\Phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ telle que, pour presque tout point x, Φ est différentiable en x avec $T(x) = \nabla\Phi(x)$.
De plus, X est caractérisée comme l'unique projection L^2 de Y sur l'ensemble $S(D)$ de toutes les applications boréliennes de D dans D conservant $\mathcal{L}$, i.e.

$$\int_D |Y(x) - X(x)|^2 d\mathcal{L}(x) < \int_D |Y(x) - \tilde{X}(x)|^2 d\mathcal{L}(x),$$

quelle que soit $\tilde{X} \in S(D)$ différent de X.
De même, $T : D \to \mathbb{R}^d$ est caractérisée comme l'unique application avec potentiel convexe telle que $\mathcal{L} \circ T^{-1} = \mathcal{L} \circ Y^{-1}$.

Il s'agit là d'une version non-linéaire de la fameuse décomposition de Helmlhotz des champs de vecteurs, qui nous dit, dans l'une de ses variantes, que tout champ $z \in L^2(D;\ \mathbb{R}^d)$ admet une unique décomposition $z = w + \nabla p$, avec w champ L^2 à divergence nulle et tangent au bord de D. De fait, si on linéarise la factorisation polaire autour de l'application identité, on retrouve bien la décomposition de Helmholtz. En effet, au moins formellement, la factorisation $Y = \nabla\Phi \circ X$, pour une application Y proche de l'identité s'écrivant $Y(x) = x + \epsilon z(x)$, donne bien $\Phi(x) = |x|^2/2 + \epsilon p(x)$, $X(x) = x + \epsilon w(x) + O(\epsilon^2)$, avec $z = \nabla p + w$, en observant que la conservation des volumes par X se traduit, pour toute fonction test f, par

$$0 = \int_D f(x + \epsilon w(x) + O(\epsilon^2))dx - \int_D f(x)dx = \int_D \nabla f(x) \cdot w(x)dx + O(\epsilon^2)$$

ce qui traduit bien, au sens faible, que w est à divergence nulle et tangent au bord de D.

Notons finalement que le terme de factorisation polaire fait référence à celle, bien connue, des matrices réelles, ce qui correspond en fait au *cas particulier* où D est la boule unité, $T(x) = Ax$, $\forall x \in D$, pour une certaine matrice A réelle $d \times d$. On a en effet alors :

$$T = \nabla\Phi \circ X, \quad \Phi(x) = \frac{1}{2}x \cdot \sqrt{AA^t}\, x,$$

et, lorsque A est non-dégénérée (i.e. inversible),

$$X(x) = Ux, \quad U = (AA^t)^{-1/2}A,$$

où U est une matrice orthogonale puisque

$$UU^t = (AA^t)^{-1/2}AA^t(AA^t)^{-1/2} = \mathbb{I}_d = U^tU,$$

$\mathbb{I}_d$ notant la matrice identité. (Notons qu'alors X fait mieux que conserver les volumes. C'est en fait une isométrie de D!)

2 Convexité cachée dans les EDP : le paradigme des équations d'Euler

Il est temps d'en venir aux EDP qui sont au centre de nos préoccupations. Il se trouve que le travail d'Euler sur les fluides incompressibles [127] de 1755/1757, va nous fournir, de plusieurs façons, des exemples remarquables de convexité caché. Il peut sembler exotique ou anecdotique d'aborder le domaine des EDP par des équations aussi particulières que celles d'Euler qui, de plus, restent si mal comprises du point de vue de l'analyse mathématique plus d'un quart de millénaire après leur introduction dans [127]. (Voir tout de même les très nombreux résultats contenus dans les livres [12, 96, 174, 189] pour n'en citer que quelques uns.) Il y a plusieurs raisons à ce parti-pris, que l'on va développer dans les prochaines sous-sections :
1) Les équations d'Euler, qui décrivent le mouvement des fluides "parfaits" (i.e. sans viscosité), aussi bien dans le cas compressible (en gros l'air) que dans le cas incompressible (en gros l'eau), constituent la première théorie des champs, où la matière est décrite en terme de champs (typiquement de densité, de vitesse et de pression). Ainsi la théorie d'Euler pour les fluides précède celles de Maxwell (Electromagnétisme), Einstein (Gravitation), Schrödinger et Dirac (Mécanique

Quantique). Elle constitue toujours la base "mécanique" de tout modèle de simulation des écoulements océaniques et atmosphériques, en météorologie comme en climatologie. (La partie "chimique" et "thermodynamique" décrivant les échanges entre soleil, terre, air, eau, sel et glace devant y être ajoutée et posant des problèmes d'une grande complexité en dehors du contexte des équations d'Euler proprement dites.) Elle est aussi à la base de la théorie de la convection si importante dans nos existences (du chauffage des logements au champ magnétique terrestre, de la cuisson dans les casseroles à la dérive des continents en passant par les tremblements de terre !). La théorie d'Euler des fluides est donc de toute première importance en termes d'applications.
2) Les équations d'Euler sont tout simplement les premières EDP jamais écrites, au moins de façon moderne. (On peut rappeler le travail légèrement antérieur de d'Alembert [2] sur les cordes vibrantes, avec sa fameuse formule permettant des les résoudre exactement, mais on pourra constater l'énorme pas en avant, rien qu'en termes de notations, que constitue la contribution d'Euler [127].)
3) Elles contiennent, de façon plus ou moins cachée, comme on va le voir brièvement un peu plus loin, toutes les EDP de base, à savoir, l'équation de la chaleur (résolue par Fourier au début du 19-ème siècle), celle des ondes et celle de Poisson et l'équation dite de transport ou encore d'advection, couvrant la classification usuelle des EDP : parabolique (chaleur), hyperbolique (ondes), elliptique (Poisson), y compris les EDO (transport).
4) Elles admettent, dans leur version incompressible, une remarquable interprétation géométrique mise en avant par Arnold [11, 12] qui les a caractérisées, dans le cas d'un fluide se mouvant sur une variété riemannienne compacte $\mathcal{M}$, comme les équations des courbes géodésiques (à vitesse constante) sur le groupe de Lie (formel) des difféomorphismes de $\mathcal{M}$ conservant l'orientation et le volume riemannien, pour la métrique L^2 sur son algèbre de Lie, laquelle n'est autre que l'espace des champs de vecteur sur $\mathcal{M}$ à divergence nulle. Dans le cas d'un fluide incompressible se déplaçant à l'intérieur du cube unité $D = [0,1]^d$, cela revient, en termes plus élémentaires, à considérer les courbes $t \in \mathbb{R} \to X_t \in SDiff(D) \subset H = L^2(D; \mathbb{R}^d)$ qui minimisent

$$\int_{t_0}^{t_1} \left|\left|\frac{dX_t}{dt}\right|\right|_H^2 dt,$$

sur tout intervalle $[t_0, t_1]$ suffisamment court, lorsque les points extrèmes X_{t_0} et X_{t_1} sont fixés. (On peut donc les voir aussi comme des applications harmoniques d'un intervalle de $\mathbb{R}$ vers $SDiff(D)$, ce qui laisse facilement imaginer des généralisations, comme des "wave maps" d'un ouvert de $\mathbb{R}^2$ vers $SDiff(D)$, qui, en fait, modélise la magnétohydrodynamique idéale des fluides incompressibles.)
Pour toutes ces raisons, on mettra souvent les équations d'Euler au premier plan, mais on verra beaucoup d'autres EDP, notamment dans le cadre des "systèmes entropiques de lois de conservation", en particulier les équations de Born-Infeld, version non-linéaire de l'Electromagnétisme de Maxwell, introduites en 1934 [34], certes beaucoup moins connues que celles d'Euler mais qui nous paraissent tout aussi remarquables et exemplaires, en se situant aussi à l'interface de

la géométrie et de la physique, à un niveau certes plus modeste que celui des équations d'Einstein mais peut-être plus accessible, au moins pour l'auteur de ces lignes.

2.1 Les équations d'Euler au coeur des EDP

Depuis qu'Euler a décrit en termes d'EDP le mouvement des fluides en 1755/1757, les EDP sont omniprésentes en mécanique, physique et géométrie. Comme on le sait bien, il n'y a pas vraiment de théorie générale des EDP (voir le cours d'Evans [128] et on pourra aussi consulter la synthèse monumentale de Taylor [226]), sauf, sans doute, dans le cas linéaire où Lars Hörmander en a fait une théorie unifiée [159]. Comme beaucoup des EDP les plus intéressantes sont quand même non-linéaires, beaucoup verront dans le travail de Jean Leray [171] en 1934 sur les équations de Navier-Stokes (une extension de celles d'Euler aux fluides visqueux) le vrai point de départ de l'analyse de l'analyse moderne des EDP où apparaissent déjà, avant la lettre, les distributions, les formulations faibles, les espaces et injections de Sobolev etc...Pourtant, il nous semble intéressant de remonter plus directement à la source, i.e. à Euler et à ses équations des fluides "parfaits" (i.e. non visqueux), ce que Richard Feynman appelait les équations de "l'eau sèche" [135]. Après les avoir exposées presqu'exactement dans les termes d'Euler [127], on va montrer, au moins formellement, comment elles contiennent déjà implicitement (au moins asymptotiquement) les EDP linéaires les plus simples: celle des ondes, celle de la chaleur, celle de Poisson, qui sont les protoypes des EDP, respectivement "hyperboliques, paraboliques et elliptiques".

"Lisez Euler, il est notre maître à tous !" (Laplace)

Voici les équations écrites par Euler en 1755/57 [127] (on utilise ici la notation ∇ alors qu'Euler écrit explicitement les dérivées partielles, avec une notation quasi-contemporaine ; on verra un peu plus loin un fac simile de l'article de 1757) :

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \partial_t v + (v \cdot \nabla) v = -\frac{1}{\rho} \nabla (p(\rho))$$

où $(\rho, p, v) \in \mathbb{R}^{1+1+3}$ sont les champs de densité, de pression et de vitesse du fluide, la pression étant supposée (par Euler) être une fonction donnée de la densité. (Euler note la densité par q plutôt que par ρ comme c'est dorénavant l'usage.) On peut encore les écrire sous forme "conservative"

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \partial_t (\rho v) + \nabla \cdot (\rho v \otimes v) = -\nabla (p(\rho))$$

et aussi "en coordonnées" (ce qui se généralise alors bien aux variétés riemanniennes) :

$$\partial_t \rho + \partial_j (\rho v^j) = 0, \quad \partial_t (\rho v_i) + \partial_j (\rho v^j v_i) = -\partial_i (p(\rho)).$$

(Dans le cas Euclidien v_i est juste une notation un peu superflue pour $\delta_{ij} v^j$ mais dans le cas Riemannien, $v_i = g_{ij} v^j$ fait impèrativement appel au tenseur métrique g.) Il est important de noter que, dans le même article, Euler considère aussi le cas incompressible

$$\partial_t v + \nabla \cdot (v \otimes v) + \nabla p = 0, \quad \nabla \cdot v = 0,$$

ou encore

$$\partial_t v_i + \partial_j (v^j v_i) = -\partial_i p, \quad \partial_i v^i = 0,$$

qui, grosso modo, correspond au cas d'une densité constante, la pression étant alors une inconnue du système, destinée à s'ajuster de sorte que le mouvement du fluide conserve les volumes. On peut d'ailleurs l'éliminer en appliquant l'opérateur divergence et on obtient alors

$$-\Delta p = (\nabla \otimes \nabla) \cdot (v \otimes v),$$

ce qui permet d'obtenir la pression en résolvant un problème de Poisson. (Le passage du cas compressible au cas incompressible a d'ailleurs été entièrement justifié mathématiquement dans les années 80 [163, 165] en faisant tendre vers l'infini la "vitesse du son", définie comme $\sqrt{\frac{dp}{d\rho}}$, ce qui suppose d'ailleurs que la pression est fonction croissante de la densité.)

Remarque.
En fait, il est important pour les applications, notamment géophysiques, de considérer aussi le cas de fluides certes incompressibles mais inhomogènes, i.e. dont la densité n'est pas uniforme. (C'est d'ailleurs le cas des écoulements océaniques.) L'inhomogénéité des fluides combinée à la force de gravité est à l'origine des mouvements de convection qui ont un énorme impact sur notre existence (du chauffage des logements au champ magnétique terrestre, de la cuisson dans les casseroles à la dérive des continents en passant par les tremblements de terre, comme nous l'avons déjà dit!). Si l'on ajoute l'effet d'un potentiel externe (comme celui de gravité) $\Phi = \Phi(x)$ -ce que d'ailleurs fait déjà Euler-, on aboutit aux équations suivantes

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \partial_t (\rho v) + \nabla \cdot (\rho v \otimes v) + \nabla p + \rho \nabla \Phi = 0, \quad \nabla \cdot v = 0,$$

où, comme ρ et v, p est une inconnue, mais qu'on peut encore éliminer grâce à une équation de Poisson

$$-\Delta p = (\nabla \otimes \nabla) \cdot (\rho v \otimes v) + \nabla \cdot (\rho \nabla \Phi).$$

Observons qu'un fluide homogène (i.e. $\rho = cte$) ne peut ressentir l'effet de Φ qui est automatiquement absorbé par la pression (ce qui fait que l'on ressent à peine la gravité sous l'eau dont la densité est presqu'égale à la nôtre).
Fin de la remarque.

Montrons maintenant, brièvement et formellement, comment les EDP de base (chaleur, ondes, Poisson, transport) sont en fait toutes déjà contenues dans celles d'Euler.

D'Euler à la chaleur par un changement quadratique du temps

Pour retrouver l'équation de la chaleur, on peut procéder au changement quadratique de la variable temps (on reviendra, beaucoup plus loin, en section 9, sur ce procédé peu conventionnel en EDP mais efficace). On part d'une solution, notée $(\tilde{\rho}, \tilde{v})$, des équations d'Euler, i.e.

$$\partial_t \tilde{\rho} + \nabla \cdot (\tilde{\rho}\tilde{v}) = 0, \quad \partial_t(\tilde{\rho}\tilde{v}) + \nabla \cdot (\tilde{\rho}\tilde{v} \otimes \tilde{v}) = -\nabla(p(\tilde{\rho}))$$

(en supposant, comme Euler, que p est fonction de la densité) et on effectue le changement de variable en temps quadratique suivant :

$$t \to \tau = t^2/2, \quad \frac{d\tau}{dt} = t, \quad (\tilde{\rho}, \tilde{v})(t,x) = (\rho(\tau,x), \frac{d\tau}{dt} v(\tau,x)),$$

(de sorte que $\tilde{v}(t,x)dt = v(\tau,x)d\tau$) On obtient alors par substitution :

$$\partial_\tau \rho + \nabla \cdot (\rho v) = 0, \;\; \rho v + 2\tau \left(\partial_\tau(\rho v) + \nabla \cdot (\rho v \otimes v)\right) = -\nabla(p(\rho)).$$

Pour les "temps petits" $\tau << 1$, on obtient une équation *asymptote* en retirant les termes en facteur de τ qui sont censés être plus petits que les autres. Il nous reste donc

$$\partial_\tau \rho + \nabla \cdot (\rho v) = 0, \quad \rho v = -\nabla(p(\rho))$$

qui, dans le cas d'un fluide "isotherme" (i.e. lorsque p est linéaire en ρ, $p = \gamma^2 \rho$, avec "vitesse du son" γ), n'est autre que la célèbrissime équation de la chaleur, résolue par Fourier au 19ème siècle (des décennies après Euler et par transformation de ... Fourier !),

$$\partial_\tau \rho = \gamma^2 \Delta \rho, \quad \Delta = \nabla \cdot \nabla.$$

Dans le cas général, on a l'équation (non-linéaire) des "milieux poreux"

$$\partial_\tau \rho = \Delta(p(\rho))$$

que nous aurons l'occasion de revoir, beaucoup plus loin en section 10.

D'Euler aux ondes par linéarisation

En écrivant que

$$(\tilde{\rho}, \tilde{v})(t,x) = (\rho^* + \epsilon\rho(t,x), \epsilon v(t,x)),$$

(avec ϵ petit et où ρ^* est une constante autour de laquelle le champ de densité fluctue) satisfait aux équations d'Euler

$$\partial_t \tilde{\rho} + \nabla \cdot (\tilde{\rho}\tilde{v}) = 0, \quad \partial_t(\tilde{\rho}\tilde{v}) + \nabla \cdot (\tilde{\rho}\tilde{v} \otimes \tilde{v}) = -\nabla(p(\tilde{\rho}))$$

(et en supposant encore que p est une fonction de ρ), on trouve

$$\partial_t \rho + \nabla \cdot ((\rho^* + \epsilon\rho)v) = 0$$

$$\partial_t((\rho^* + \epsilon\rho)v) + \nabla \cdot ((\rho^* + \epsilon\rho)\epsilon v \otimes v) = -\nabla \left(\frac{p(\rho^* + \epsilon\rho) - p(\rho^*)}{\epsilon} \right)$$

Dans le régime $\epsilon << 1$, pour de "petites fluctuations de densité et de vitesse", on obtient encore une équation asymptote en éliminant les termes petits en ϵ. (tout en utilisant $p(\rho^* + \epsilon\rho) = p(\rho^*) + \epsilon p'(\rho^*)\rho + O(\epsilon^2)$). Il nous reste donc

$$\partial_t \rho + \rho^* \nabla \cdot v = 0, \quad \rho^* \partial_t v + p'(\rho^*) \nabla \rho = 0$$

qui n'est autre que la fameuse équation des ondes (presqu'aussi fameuse que celle de la chaleur et qu'avait résolue en une dimension d'espace d'Alembert quelques années avant le travail d'Euler sur les fluides [2]) :

$$\partial_{tt}^2 \rho = \gamma^2 \Delta \rho$$

(après élimination de v), avec "vitesse du son" $\gamma = \sqrt{p'(\rho^*)}$.

Equations d'Euler dans le plan comme *couplage* d'équations linéaires

Dans le cas incompressible, où $\nabla \cdot v = 0$, et dans le plan (et plus généralement dans un domaine simplement connexe du plan), on peut écrire

$$v = (-\partial_2 \psi, \partial_1 \psi)$$

pour une certaine fonction scalaire $\psi = \psi(t, x)$ (dite “fonction de courant”). En posant $\omega = \partial_2 v_1 - \partial_1 v_2$, on observe que, d'une part

$$-\Delta \psi = \omega$$

et d'autre part (par un petit calcul élémentaire)

$$\partial_t \omega + (v \cdot \nabla)\omega = 0.$$

Dans ce cas, les équations d'Euler ne sont que le couplage *non trivial* de deux équations linéaires très simples :
1) L'équation de Poisson, prototype des équations elliptiques,

$$-\Delta \psi = \omega$$

posée en ψ avec ω comme donnée ;
2) L'équation de transport (ou d'advection), qui n'est qu'une EDO déguisée (comme on va le voir dans un instant),

$$\partial_t \omega + (v \cdot \nabla)\omega = 0.$$

posée en ω avec $v = (-\partial_2 \psi, \partial_1 \psi)$ comme donnée.

Les équations d'Euler et les EDO

En intégrant le champ de vitesse v du fluide, on récupère les trajectoires des particules fluides :

$$\frac{dX_t}{dt}(a) = v(t, X_t(a)),$$

où a est l'étiquette de la particule et $X_t(a)$ sa position au temps t. (Il est usuel mais nullement nécessaire d'étiqueter la particule par sa position initiale, i.e. $X_0(a) = a$.) Par dérivation composée, on voit aisément que l'équation d'Euler

$$\partial_t v + (v \cdot \nabla)v = -\frac{\nabla p}{\rho}$$

n'est rien d'autre que la traduction de l'EDO

$$\frac{d^2 X_t}{dt^2}(a) = -(\frac{\nabla p}{\rho})(t, X_t(a))$$

et, même, dans le cas des fluides incompressibles homogènes :

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

En fait, dans son article [127], Euler *part* de cette EDO, après avoir expliqué en détail le concept de champ de pression, pour *obtenir* les équations Euler précisément en introduisant le concept de champ de vitesse. (Ce point essentiel est assez souvent omis dans la littérature.) Le lien avec les EDO est encore plus frappant dans le cas bidimensionnel incompressible. En effet, l'équation du tourbillon que l'on vient de voir et qui peut s'écrire

$$\partial_t \omega + (v \cdot \nabla)\omega = 0$$

signifie juste que $\Omega(t, a) = \omega(t, X_t(a))$ est indépendant de t. En effet, l'équation de transport du tourbillon se réduit à l'EDO triviale :

$$\frac{d\Omega}{dt} = 0.$$

Quelques mots sur l'analyse des équations d'Euler

Jusqu'à maintenant, on n'a pas du tout abordé les équations d'Euler du point de vue de l'analyse. On a suivi en cela l'exemple d'Euler lui-même qui termine son article par la conclusion prophétique :
"Tout ce que la théorie des fluides renferme est contenu dans les deux équations rapportées ci-dessus, de sorte que ce ne sont pas les principes de Mécanique qui nous manquent dans la poursuite de ces recherches, mais uniquement l'Analyse, qui n'est pas encore assez cultivée, pour ce dessein."
Un quart de millénaire plus tard, les progrès restent en effet limités. Ainsi, les équations d'Euler sont les plus anciennes (ou presque) des EDP mais constituent toujours un défi majeur pour les analystes. Voici quelques résulats (cf.

[95, 174, 188, 189]...).
Commençons par le cas incompressible homogène.

On sait (en plaçant sur le cube périodique $D = \mathbb{T}^d$, pour simplifier) que :
1) une solution classique existe toujours en temps petit, et est unique, dès que le champ de vitesse initialement donné v_0 est assez régulier (i.e. de dérivées Hölder continues) (Wolibner 1933 [234]);
2) dans le plan ($d = 2$), une unique solution globale existe (en un sens à préciser) dès que le tourbillon initial ω_0 (i.e. le rotationnel de v_0) est essentiellement borné sur D (Yudovich 1963 [236]) et la régularité des lignes de niveau du tourbillon est propagée sans limitation de temps (Chemin 1991 [93]); si ω_0 est somme d'une fonction L^1 et d'une mesure positive bornée, l'existence globale d'une solution faible est assurée, mais pas l'unicité (Delort 1991 [110]);
3) des solutions faibles $v \in L^2$ au sens des distributions existent toujours globalement en temps dès que $v_0 \in L^2(D;\ \mathbb{R}^d)$, mais malheureusement elles sont en nombre infini! (Wiedemann 2011 [232], comme corollaire des travaux de De Lellis et Székelyhidi -par "intégration convexe"- sur lesquels on va bientôt revenir [107, 108]); d'autre part, par les mêmes méthodes, on montre l'existence de solutions faibles $v(t, x)$ qui sont Hölder continues d'exposant inférieur à $1/3$ en x pour lesquelles l'énergie cinétique n'est pas conservée (résolution de la "conjecture d'Onsager" par Isett 2018 [160, 86]), alors que pour un exposant supérieur à $1/3$, l'énergie est connue pour être conservée [98, 131].
4) des solutions généralisées globales, dites "dissipatives", existent toujours dès que $v_0 \in L^2(D;\ \mathbb{R}^d)$; elles ne sont pas forcément solutions faibles mais ont l'avantage de ne pouvoir différer d'une éventuelle solution classique (Lions 1996 [174]); nous reviendrons sur ce concept qui est typique de la méthode convexe d'entropie relative et qu'on utlisera souvent dans ce cours;
5) du point de vue géométrique, le flot géodésique sur le groupe $SDiff(D)$ est bien défini de façon classique, mais seulement dans un petit voisinage de l'identité et pour une topologie très fine (de type Sobolev) (Ebin-Marsden 1970 [121]). En revanche pour $d = 3$, on peut trouver de nombreux difféomorphismes triviaux dans la 3ème dimension, i.e. de la forme $h(x_1, x_2, x_3) = (H(x_1, x_2), x_3)$, pour lesquels il ne peut y avoir de géodésiques minimisantes -au sens classique- les reliant à l'identité (Shnirelman 1985 [219]). Dans ce cas, on peut pourtant montrer, par une technique de convexification qu'on va discuter bientôt, l'existence d'un unique gradient de pression ∇p, entièrement déterminé par H, telles que toutes les géodésiques minimisantes approchées sont approximativement "accélérées" par $-\nabla p$ (Brenier 1999 [48]), le résulat s'étendant en fait à tous les h de $SDiff(D)$ (Ambrosio-Figalli 2008 [5]).

Dans le cas compressible, la situation est encore pire. On a grosso modo l'équivalent des quatre premiers résultats (mais, pour le second, seulement avec $d = 1$ et avec des restrictions de taille sur les données initiales [82, 31]).
Bien entendu, il y a beaucoup d'autres résultats importants mais les cinq que l'on vient d'énoncer sont très représentatifs de notre point de vue.

Après cette revue dés équations d'Euler, on va voir comment la convexité se cache, dans un premier temps dans leur version incompressible, autant du point de vue géométrique à la Arnold que du point plus classique des EDP et on verra comment le point de vue "trajectoriel" permet d'avoir un éclairage commun. Dans tous les cas, on verra émerger des théorèmes généralisant au cadre des EDP les résultats de convexité cachée décrits dans le chapitre introductif.

2.2 Convexité cachée dans les équations d'Euler : le point de vue géométrique

Comme on l'a vu plus haut, une définition plus géométrique (suivant Arnold [11]) des équations d'Euler revient à en chercher les solutions comme des courbes $t \in \mathbb{R} \to X_t \in SDiff(D) \subset H = L^2(D;\mathbb{R}^d)$ qui minimisent

$$\int_{t_0}^{t_1} ||\frac{dX_t}{dt}||_H^2 dt,$$

sur tout intervalle $[t_0, t_1]$ suffisamment court, lorsque les points extrèmes X_{t_0} et X_{t_1} sont fixés. Reprenons cela en termes d'un concept très en vogue à la fin du 18ème siècle, celui de "principe de (la) moindre action".

(La paternité du principe de la moindre action était alors attribuée à Maupertuis, président de l'Académie Royale des Sciences de Berlin, en particulier par Euler, membre de cette même Académie. Le mathématicien Koenig prétendit disposer d'une lettre prouvant qu'en fait l'idée était bien antérieure et provenait de Leibniz, ce qui suscita une violente attaque de l'Académie, menée par Euler lui-même, accusant Koenig de fraude. Il s'ensuivit une virulante polémique, couronnée par un pamphlet de Voltaire se moquant de Maupertuis -sous le sobriquet de Dr. Akakia- qui remporta un grand succès à Paris mais que le roi Frédéric II de Prusse s'empressa de faire brûler dans son royaume.)

Les équations d'Euler des fluides incompressibles obéissent en effet au "beau principe de (la) moindre action" (pour reprendre l'expression même d'Euler [127]), l'espace de configuration étant le groupe $SDiff(D)$ des difféomorphismes conservant volume et orientation et l'énergie cinétique étant définie par le carré de la norme L^2 de la dérivée de la transformation X_t par rapport au temps t.

Arrêtons-nous sur ce dernier point pour un instant. Grâce à l'interprétation géométrique des équations d'Euler, posées dans un domaine compact $D \subset \mathbb{R}^d$, par exemple le cube unité, on est renvoyé à la discussion du premier chapitre concernant la compactification-convexification du groupe $SDiff(D)$ des difféomorphismes de D conservant volume et orientation par l'ensemble convexe compact $DS(D)$ des mesures bistochastiques sur D. Il est tentant d'adapter le principe de moindre action en le levant du niveau de $SDiff(D)$, qui n'a le mérite ni d'être compact, ni même fermé dans le Hilbert $H = L^2(D;\mathbb{R}^d)$, au

niveau du convexe compact $DS(D)$. C'est ce qui va nous occuper pendant toute une partie de ce cours. L'idée de départ est fort simple. Esquissons-là dès à présent, sachant que nous la reprendrons et la développerons plus avant dans le cours.
A une courbe $t \to X_t \in SDiff(D)$ on attachera la courbe de mesure bistochastique associée $t \to c_t \in DS(D)$, en posant

$$\int_{D^2} f(x,a)dc_t(x,a) = \int_D f(X_t(a),a)da, \quad \forall f \in C^0(D^2)$$

comme on l'a déjà fait plus haut. On écrira aussi bien :

$$dc_t(x,a) = \delta(x - X_t(a))da.$$

Cela est hélas insuffisant pour décrire la dynamique. L'idée clé (avatar de celle de "courant", classique en théorie géométrique de la mesure [132, 194, 235]) est de lui associer une courbe

$$t \to q_t \in \left(C^0(D^2;\mathbb{R}^d)\right)'$$

de mesures à valeurs vectorielles (i.e. dans le dual des fonctions continues sur D^2 à valeurs dans $\mathbb{R}^d$), définies par

$$\int_{D^2} f(x,a) \cdot dq_t(x,a) = \int_D \frac{dX_t}{dt}(a) \cdot f(X_t(a),a)da, \quad \forall f \in C^0(D^2;\mathbb{R}^d)$$

où $\frac{dX_t}{dt}(a)$ n'est rien d'autre qu'une notation pour la dérivée partielle $\partial_t X_t(a)$, l'application T étant vue comme fonction de t et a. On pourra aussi écrire, plus brièvement

$$dq_t(x,a) = \frac{dX_t}{dt}(a)\delta(x - X_t(a))da.$$

Ces deux mesures sont liées l'une à l'autre, d'abord par une relation d'ordre, à savoir que q_t est par construction absolument continue par rapport à c_t et admet donc une dérivée de Radon-Nikodym à valeurs vectorielles, qu'on notera $(x,a) \to v_t(x,a) \in \mathbb{R}^d$, de sorte qu'on pourra écrire :

$$dq_t(x,a) = v_t(x,a)dc_t(x,a).$$

Ensuite, du moment que X_t est différentiable en t (en fait, techniquement, il suffit que pour presque tout $a \in D$, $t \to X_t(a)$ soit absolument continue, i.e. primitive en t d'une fonction Lebesgue intégrable), les deux mesures sont différentiellement liées par l'équation

$$\partial_t c_t + \nabla_x \cdot q_t = 0,$$

satisfaite au sens des distributions. En effet, pour toute fonction test $f = f(x,a)$ définie sur $D \times D$, on a

$$\frac{d}{dt}\int_{D^2} f(x,a)dc_t(x,a) = \frac{d}{dt}\int_D f(X_t(a),a)da$$

$$= \int_D (\nabla_x f)(X_t(a), a) \cdot \frac{dX_t}{dt}(a) da = \int_{D^2} f(x,a) dq_t(x,a).$$

D'autre part, on peut se convaincre que l'énergie cinétique peut se réécrire simplement en termes de c et v seulement :

$$\frac{1}{2}||\frac{dX_t}{dt}||_H^2 = \frac{1}{2}\int_{D^2} |v_t(x,a)|^2 dc_t(x,a).$$

Pour le prouver, le mieux est de procéder par dualité, en partant de la droite. On observe d'abord (par "complétion des carrés") que

$$\frac{1}{2}\int_{D^2} |v_t(x,a)|^2 dc_t(x,a) =$$

$$\sup\{\int_{D^2} \left(-\frac{1}{2}|f(x,a)|^2 + f(x,a)\cdot v_t(x,a)\right) dc_t(x,a); \quad f \in C^0(D^2;\ \mathbb{R}^d)\}$$

(selon la théorie de la mesure, par densité des fonctions continues dans l'espace des fonctions de carré sommable par rapport à la mesure c_t)

$$= \sup\{\int_{D^2} \left(-\frac{1}{2}|f(x,a)|^2 dc_t(x,a) + f(x,a)\cdot dq_t(x,a)\right); \quad f \in C^0(D^2;\ \mathbb{R}^d)\}.$$

$$= \sup\{\int_{D} \left(-\frac{1}{2}|f(X_t(a),a)|^2 + f(X_t(a),a)\cdot \frac{dX_t}{dt}(a)\right) da; \quad f \in C^0(D^2;\ \mathbb{R}^d)\}.$$

$$= \frac{1}{2}\int_D |\frac{dX_t}{dt}(a)|^2 da$$

(de nouveau par complétion des carrés, en utilisant que $a \to X_t(a)$ est bijective puisque X_t appartient à $SDiff(D)$). Cette série d'égalités est d'autant plus intéressante qu'elle nous donne une expression *convexe* de l'énergie cinétique en fonction des mesures c et m, sous la forme intermédiaire

$$\sup\{\int_{D^2} \left(-\frac{1}{2}|f(x,a)|^2 dc_t(x,a) + f(x,a)\cdot dq_t(x,a)\right); \quad f \in C^0(D^2;\ \mathbb{R}^d)\},$$

qu'on vient d'obtenir. On peut même améliorer cette formule en écrivant, pour un couple quelconque de mesures $(c,q) \in \left(C^0(D^2;\ \mathbb{R}\times\mathbb{R}^d)\right)'$

$$K(c,q) = \sup\{\int_{D^2} A(x,a)dc(x,a) + B(x,a)\cdot dq(x,a);$$

$$(A,B) \in C^0(D^2;\ \mathbb{R}\times\mathbb{R}^d) \text{ s.t. } 2A + |B|^2 \le 0\},$$

qui a l'avantage de définir une fonction convexe s.c.i à valeurs dans $]-\infty,+\infty]$ sans aucune restriction sur $(c,q) \in \left(C^0(D^2;\ \mathbb{R}\times\mathbb{R}^d)\right)'$, même pas que $c \ge 0$. On peut en effet montrer que $K(c,q)$ prend toujours la valeur $+\infty$ sauf si :

$c \geq 0$, q est absolument continue par rapport à c, de dérivée de Radon-Nikodym v de carré intégrable en c, auquel cas elle prend la valeur

$$\frac{1}{2}\int_{D^2} |v(x,a)|^2 dc(x,a).$$

(On reviendra plus loin sur ces calculs, qu'on pourra admettre pour l'instant, bien qu'il s'agisse d'un exercice élémentaire de théorie de la mesure.) Ainsi on est amené à retranscrire le principe de moindre action entièrement en termes des mesures (c_t, q_t) en leur demandant de minimiser, pour tout intervalle de temps $[t_0, t_1]$ assez petit

$$\int_{t_0}^{t_1} K(c_t, q_t)dt,$$

sous les contraintes que c_t est bistochastique, i.e. $c_t \in DS(D)$, que l'EDP linéaire

$$\partial_t c_t + \nabla_x \cdot q_t = 0$$

soit satisfaite et qu'enfin les valeurs aux limites c_{t_0} et c_{t_1} soient fixées dans $DS(D)$. La nouveauté de cette formulation est que l'on n'impose plus à c et q de dériver d'une courbe $t \to X_t \in SDiff(D)$. Il s'agit donc d'une version relaxée de la formulation initiale, qui présente l'immense avantage d'être entièrement convexe.
En termes plus géométriques, on peut aussi voir dans ce problème de minimisation sous contraintes, lorsque t_0 et t_1 sont fixés, et même normalisés comme $t_0 = 0$, $t_1 = 1$ avec c_0 et c_1 données dans $DS(D)$ la version convexifiée du problème de "géodésiques minimisantes" entre deux points de l'espace de configuration $SDiff(D)$. L'étude détaillée de ce problème sera menée dans le chapitre 4 de ce cours. En voici une première synthèse des resultats (suivant [48] et poursuivis dans [5, 6, 62, 15]).
To make notations lighter, we will write $c(t,x,a), q(t,x,a), v(t,x,a)$ instead of $c_t(x,a), q_t(x,a), v_t(x,a)$ as we did before and we will use shorter notations for integrals, such as

$$\int_{x,a} f(x,a)c(t,x,a)$$

rather than

$$\int_{D^2} f(x,a)dc_t(x,a).$$

Theorem 10. *Let D be the periodic cube $D = \mathbb{T}^d$. Given any data c_0 and c_1 in the convex compact set of all doubly stochastic measure on D, the relaxed minimizing geodesic problem always admits at least one solution (c, cv) and there is a unique pressure gradient $(t,x) \in]0,1[\times D \to \nabla p(t,x) \in \mathbb{R}$, depending only on c_0 and c_1 such that*

$$-\Delta p(t,x) = \nabla_x \otimes \nabla_x \cdot \int_a (cv \otimes v)_t(x,a)$$

whatever solution (c, cv) is.
In addition, ∇p has some limited regularity: it is locally square integrable in time with values in the space of bounded measures on $D = \mathbb{T}^d$. As a consequence, we get the relaxed Euler equation in form

$$(\partial_t(cv) + \nabla_x \cdot (cv \otimes v))(t, x, a) = -\underline{c}(t, x, a)\nabla_x p(t, x),$$

where $\underline{c}$ is a suitable extension of c to the (possibly) singular set of $\nabla_x p$ (which is a priori just a Borel measure).
Moreover, each optimal solution (c, cv) can be weakly- approximated by a family of smooth curves $t \in [0, T] \to T_t^\epsilon \in SDiff(D)$, in the sense that, denoting*

$$v^\epsilon = \frac{dT_t^\epsilon}{dt} \circ (T_t^\epsilon)^{-1},$$

the corresponding measures

$$(1, \frac{dT_t^\epsilon}{dt}(a))\delta(x - T_t^\epsilon(a))$$

weakly- converge to $(c, cv)(t, x, a)$ and without gap of energy, in the sense*

$$\int_0^1 \int_D |v^\epsilon(t, x)|^2 dxdt \to \int_0^1 \int_{x,a} (c|v|^2)(t, x, a)dt.$$

Finally, the v^ϵ are almost solutions to the Euler equations in the sense that

$$\partial_t v^\epsilon + \nabla \cdot (v^\epsilon \otimes v^\epsilon) \to -\nabla p,$$

where ∇p is the unique pressure gradient attached to the data (c_0, c_1).

Let us epmhasise that it is very surprising that the pressure gradient is uniquely determined by the data. Indeed, let us consider, as Arnold did in his founding paper [11], the finite dimensional counterpart of the Euler model of incompressible fluids, namely the model of rigid bodies, where the finite dimensional Lie group $SO(3)$ substitutes for $SDiff(D)$, and a non-degenerate quadratic form (corresponding to the matrix of inertia of the rigid body) substitute for the L^2 metric. Then the geodesic curves precisely describe the motion of a perfect rigid body moving in vaccuum (without external forces). There is also a substitute for the pressure gradient, which turns out to be a 3×3 symmetric time dependent matrix which is attached to each geodesic, and acts in order to preserve the rigidity of the body. Then one can find examples of two minimizing geodesics having the same end-points for which these matrices are not the same [62]. As a matter of fact, the uniqueness of the pressure gradient is, in our opinion, a striking manifestation of "hidden convexity" due to to the infinite dimension of $SDiff(D)$ and the convexity of its weak completion $DS(D)$. So, in some sense, we have a rather sophisticated avatar of Theorem 2 (where, in a Hilbert, the unit ball is the right weak completion of the unit ball if only if its dimension is infinite).

There is certainly some room to improve the results we have just mentioned. In particular, it would be very useful to know the precise regularity of the pressure field. There is some evidence [62] that the pressure $p(t,x)$ should be, locally in time in $]0,1[$, semi-concave in x, and not more in general, which means that the derivatives in x of p should be Borel measures up to *second* order and not only to first order as in the Theorem!
To conclude this sub-section, let us just us mention a striking addtional property: the "Boltzmann entropy"

$$\int_{x,a} (c \log c)(t,x,a)$$

is convex in t along every generalized minimizing geodesic. This has been conjectured in [54] and proven first by Lavenant [169] (with some restrictions) and then by Baradat-Monsaingeon [18]. In our opinion, this convexity might be an indication that $SDiff(D)$ has, in some suitable sense, a nonnegative Ricci curvature (in the spirit of Lott-Sturm-Villani [183, 223]). This is another striking manifestation of “hidden convexity”, since, in the classical framework, the measures $c(t,x,a)$ are delta measures and their Boltzmann entropy is *always* infinite!

2.3 Equations d’Euler et convexité : le point de vue Eulérien

Dans ce paragraphe, nous nous limitons encore aux équations d’Euler des fluides *incompressibles*, pour lesquelles nous venons d’utiliser la formulation géométrique d’Arnold en termes de géodésiques le long du groupe $SDiff(D)$ (des difféomorphismes d’un domaine compact $D \subset \mathbb{R}^d$ qui conservent volume et orientation) pour aboutir à la "convexification" du problème des géodésiques minimisantes, posée sur l’ensemble convexe compact $DS(D)$ (des mesures bistochastiques sur D). Nous revenons maintenant au point de vue plus classique des EDP où, rappelons-le, les équations d’Euler s’écrivent simplement

$$\partial_t v + \nabla \cdot (v \otimes v) + \nabla p = 0, \quad \nabla \cdot v = 0,$$

et ou on se placera, pour simplifier, sur le cube périodique $D = \mathbb{T}^d$. Une apparition frappante de la convexité résulte des célebres travaux De Lellis et Székelyhidi [107, 108, 109], faisant appel au concepts d’inclusions différentielles et d’intégration convexe, auquel on peut attacher les noms de Gromov, Nash et Tartar [156, 197, 225]. Ils ont fait suite aux travaux de Constantin-E-Titi, Eyink, Scheffer, Shnirelman, aussi bien sur la "conjecture d’Onsager" [98, 131] que sur l’existence de solutions faibles non nulles ayant un support compact en temps (!) [211, 221]. Ils ont suscité depuis de nombreux travaux (par exemple [86, 108, 109, 160] parmi beaucoup d’autres). Un élément clé, bien qu’élémentaire, est le concept *convexe* de "sous-solution". Plus précisémet, on appelle sous-solution des équations d’Euler un couple (V, M) tel que,
i) d’une part

$$\partial_t V + \nabla \cdot M + \nabla p = 0, \quad \nabla \cdot V = 0,$$

ou encore, en coordonnées

$$\partial_t V^i + \partial_j M^{ij} + \partial^i p = 0,$$

a lieu au sens des distributions pour une certaine fonction scalaire p (la "pression");
ii) $M \geq V \otimes V$ a lieu au sens des matrices symétriques et des distributions.
On voit immédiatement que, pour devenir une solution "faible", au sens traditionnel, il suffit pour (V, M) de "saturer" l'inégalité $M \geq V \otimes V$ en $M = V \otimes V$. Notons que, du point de vue des espaces fonctionnels, le concept de sous-solution ne demande a priori pas beaucoup de régularité pour (V, M). Si on se place, pour simplifier, sur $Q = [0, T] \times D$ avec $D = \mathbb{T}^d$, il fait sens dès que $V \in L^2(Q; \mathbb{R}^d)$ et dès que M est une mesure de Borel bornée prenant ses valeurs dans le cône des matrices symétriques semi-définies positives. On peut ajouter comme condition initiale, au sens des distributions, un champ à divergence nulle de carré intégrable V_0, en imposant

$$\int_Q \partial_t A_i(t,x) V^i(t,x) dtdx + \partial_j A_i(t,x) M^{ij}(dtdx) + \int_D V_0^i(x) A_i(0,x) dx = 0,$$

pour toute champ lisse à divergence nulle A qui s'annule en $t = T$. (Attention, comme a priori M n'est ici qu'une mesure, $V(t,x)$ n'est pas mieux qu'à variation bornée en t et peut donc ne pas avoir V_0 comme limite en $t \downarrow 0$. On retrouvera ce problème un peu subtil beaucoup plus loin dans le cours en section 10 -et on le connait d'ailleurs fort bien, dans un domaine voisin des EDP, lorsqu'on traite les conditions aux limites des lois de conservation hyperboliques, à la suite du travail de Bardos, Le Roux et Nédelec [20].) Insistons sur le fait que, pour une donnée initiale v_0 fixée, l'ensemble des sous-solutions est bien convexe!
Quand l'inégalité est stricte: $M > V \otimes V$, on parle de sous-solution stricte. Inversement, les solutions faibles correspondent au cas où l'inégalité est partout saturée. On se retrouve dans une situation assez analogue à celle du Théorème 2 de l'introduction où on compare la sphère unité à la boule unité. Et, de fait, une des conséquences de la théorie de De Lellis et Székelyhidi [107] est le résultat remarquable suivant, tiré de [109] :

Theorem 11. *Soit (V, M) une sous-solution stricte et lisse des équations d'Euler sur $[0, T] \times \mathbb{T}^d$, alors il existe une suite de solutions faibles v_n des équations d'Euler, continues (et qu'on peut même supposer Hölderiennes d'exposant suffisamment modéré, mais pas supérieur à un tiers), qui converge vers (V, M) au sens que $v_n - V$ et $v_n \otimes v_n - M$ convergent faiblement-* vers zéro dans $L^\infty(\mathbb{T}^d)$, uniformément en t. On peut de plus supposer que, pour tout $t \in [0, T]$,*

$$\int_{T^d} (v_n \otimes v_n)(t,x) dx = \int_{T^d} M(t,x) dx.$$

Il s'agit d'un résultat hautement non trivial qui nécessite un travail d'analyse considérable et on se contentera ici de renvoyer le lecteur intéressé aux travaux de De Lellis et Székelhydi [107, 109].

On peut voir ce résultat comme une version très sophistiquée du théorème 2 exposé dans l'introduction du cours (les sous-solutions strictes jouant le rôle des points intérieurs à la boule et les solutions faibles celui des points de la sphère).

2.4 Various formulations of the Euler equations

The trajectorial viewpoint

It is very instructive to look at the Euler equations of incompressible homogeneous fluids at the level of trajectories (in so-called "Lagrangian coordinates"). As we already saw, they just read

$$\frac{d^2X_t}{dt^2}(a) = -(\nabla p)(t, X_t(a)),\ \forall t\ \mathcal{L} \circ X_t^{-1} = \mathcal{L} = \text{Lebesgue},$$

where a denotes the label of a typical fluid particle and $X_t(a)$ its location in the domain D at time t. (Let us recall that this is the very starting point of Euler's paper [127]! The main point of his paper was *precisely* the derivation of the *Eulerian* equations that have become so popular that many people ignore their origin which is definitely on the trajectorial -or so-called "Lagrangian" side.) Indeed, Euler postulated the existence of a vector field $v = v(t, x)$, the so-called "Eulerian velocity field" such that

$$v(t, X_t(a)) = \frac{dX_t}{dt}(a).$$

Thus, by the chain rule and assuming X_t to be one-to-one in D, one easily gets, as Euler did,

$$(\partial_t + v \cdot \nabla)v + \nabla p = 0, \quad \nabla \cdot v = 0,$$

which is the "non-conservative" form of the Euler equations, usually written as

$$\partial_t + \nabla \cdot (v \otimes v) + \nabla p = 0, \quad \nabla \cdot v = 0.$$

Very much as we did in the geometrical framework, let us introduce the "mixed Eulerian-Lagrangian" measures

$$c(t, x, a) = \delta(x - X_t(a)), \quad q(t, x, a) = \frac{dX_t}{dt}(a)\delta(x - X_t(a)),$$

which are defined on the space $[0, T] \times D \times \mathcal{A}$, where $\mathcal{A}$ is the space of "fluid particle labels". (It is customary, but in no way necessary, as will be seen later on, to define $\mathcal{A}$ as D itself, with the convention that a is nothing but the "initial position" $X_0(a)$ of the particle with label a. We just assume $\mathcal{A}$ to be a compact metric space with a probability measure on it, denoted by da for simplicity.) As observed before, from its very definition, q is absolutely continuous with respect to c and therefore it makes sense to consider its Radon-Nikodym derivative that will be denoted by $v = v(t, x, a)$, so that we will write

$$q(t, x, a) = v(t, x, a)c(t, x, a) = (cv)(t, x, a).$$

With such notations, we may write

$$\int_{x,a} f(x,a)c(t,x,a) = \int_{\mathcal{A}} f(X_t(a),a)da,$$

$$\int_{x,a} f(x,a)q(t,x,a) = \int_{x,a} f(x,a)(cv)(t,x,a) = \int_{\mathcal{A}} \frac{dX_t}{dt}(a)f(X_t(a),a)da,$$

for all continuous function f on $D\times\mathcal{A}$ and all $t \in [0,T]$. By standard differential calculus, we can get a consistent system of PDEs for (c,v) together with ∇p. The following computations are perfectly rigorous as long as $\nabla p(t,x)$ is sufficiently smooth, say Lipschitz continuous in $x \in D$ with a Lipschitz constant integrable in $t \in [0,T]$:

Proposition 12. *Let $\nabla p(t,x)$ be sufficiently smooth, say Lipschitz continuous in x, for $(t,x) \in [0,T] \times D$, where $D = \mathbb{T}^d$. Assume that $(X_t,\ t \in [0,T])$ is a family of measure-preserving maps in the sense that*

$$\int_{\mathcal{A}} f(X_t(a))da = \int_D f(x)dx,$$

for all $f \in C(D)$ and all $t \in [0,T]$. Further assume, that

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

holds true for all $a \in \mathcal{A}$ and $t \in [0,T]$.
Then the measures $(c, q = cv)$, associated with $(X_t,\ t \in [0,T])$ through

$$\int_{x,a} f(x,a)c(t,x,a) = \int_{\mathcal{A}} f(X_t(a),a)da,$$

$$\int_{x,a} f(x,a)q(t,x,a) = \int_{x,a} f(x,a)(cv)(t,x,a) = \int_{\mathcal{A}} \frac{dX_t}{dt}(a)f(X_t(a),a)da,$$

for all continuous function f on $D \times \mathcal{A}$ and all $t \in [0,T]$, satisfy the following set of equations

$$\int_a c(t,x,a) = 1, \quad \partial_t c(t,x,a) + \nabla_x \cdot (cv(t,x,a)) = 0,$$

$$(\partial_t(cv) + \nabla_x \cdot (cv \otimes v))(t,x,a) = -c(t,x,a)\nabla_x p(t,x).$$

In addition, by integrating these equations in a, we also have

$$\nabla \cdot \int_a (cv)(t,x,a) = 0, \quad -\Delta_x p(t,x) = \nabla_x \otimes \nabla_x \cdot \int_a (cv \otimes v)(t,x,a).$$

Proof:

First, since X_t is volume-preserving, we get for all test functions $f = f(x)$:

$$\int_{x,a} f(x)c(t,x,a) = \int_D f(X_t(a))da = \int_D f(x)dx.$$

Thus: $\int_a c(t,x,a) = 1$ immediately follows. Next,

$$\frac{d}{dt}\int_{x,a} f(x,a)c(t,x,a) = \frac{d}{dt}\int f(X_t(a),a)da = \int (\nabla_x f)(X_t(a),a)\cdot\frac{dX_t}{dt}(a)da$$

$= \int_{x,a} \nabla_x f(x,a)\cdot(cv)(t,x,a)$, for all test functions $f = f(x,a)$. Similarly:

$$\frac{d}{dt}\int_{x,a} f(x,a)(cv)(t,x,a) = \frac{d}{dt}\int f(X_t(a),a)\frac{dX_t}{dt}(a)da$$

$$= \int (\nabla_x f)(X_t(a),a)\cdot(\frac{dX_t}{dt}\otimes\frac{dX_t}{dt})(a)da - \int f(X_t(a),a)(\nabla_x p)(t,X_t(a))da$$

$$= \int_{x,a} \nabla_x f(x,a)\cdot(cv\otimes v)(t,x,a) - f(x,a)c(t,x,a)\nabla_x p(t,x).$$

as announced. Finally,

$$-\Delta p(t,x) = \nabla_x\otimes\nabla_x\cdot\int_a (cv\otimes v)(t,x,a).$$

just follows from

$$\int_a c(t,x,a) = 1, \quad \nabla\cdot\int_a (cv)(t,x,a) = 0.$$

End of proof.
So, the relaxed equations we have derived by pure differential calculus from the *original* Euler's model, written in terms of trajectories rather than in terms of "eulerian" fields, are nothing but the optimality conditions we have stated for the relaxed version of the minimizing geodesic, as just seen in subsection 2.3. Let us recall that this relaxed problem reads, in short,

$$\inf\{\int_0^1 dt\int_{x,a} c|v|^2\ ;\ \partial_t c + \nabla_x\cdot(cv) = 0, \quad \int_a c = 1\}$$

with $c(t,x,a)$ prescribed at $t = 0$ and $t = 1$, and is *convex* in (c, cv).

Remark.
As a matter of fact (we will go back to that later on), the optimality conditions contain an extra condition: $\nabla_x \times v(t,x,a) = 0$, that has a variational interpretation in terms of principle of least action (in relationship with Noether's

celebrated invariance theorem) and says that the velocity field $v(\cdot, \cdot, a)$ attached to the label a is curl-free. This does not contradict that the averaged velocity

$$\int_a (cv)(t, x, a)$$

is divergence-free. As a matter of fact, this provides a striking example of a *macroscopic* divergence-free vector field that can written as a linear superposition of a family of curl-free vector fields.
End of remark.

Relaxed solutions versus sub-solutions

By averaging out the relaxed solutions of the Euler equations, we immediately get some sub-solutions of the Euler equations, just by setting

$$V(t, x) = \int_a (cv)(t, x, a), \quad M(t, x) = \int_a (cv \otimes v)(t, x, a).$$

Indeed,

$$\partial_t V + \nabla \cdot M + \nabla p = 0, \quad \nabla \cdot v = 0,$$

just follow from the relaxed equations

$$\partial_t c(t, x, a) + \nabla_x \cdot (cv)(t, x, a) = 0, \quad \int_a c(t, x, a) = 1,$$

$$(\partial_t (cv)(t, x, a) + \nabla_x \cdot (cv \otimes v))(t, x, a) = -\underline{c}(t, x, a) \nabla_x p(t, x),$$

after integration in a and,

$$M \geq V \otimes V$$

is just a consequence of Jensen's inequality since $\int_a c(t, x, a) = 1$. Notice that these sub-solutions have no reason to be strict and, therefore, the De Lellis-Székelyhidi Theorem 11 a priori does not apply to them.

Relaxed versus kinetic solutions

There is a parallel formulation of the relaxed equation, of Vlasov or "kinetic" type, involving the "kinetic" "phase-density"

$$f(t, x, \xi) = \int_a \delta(\xi - v(t, x, a)) c(t, x, a), \quad (x, \xi) \in \mathbb{T}^d \times \mathbb{R}^d.$$

(Here f is a traditional notation in kinetic theory for the phase density and the letter f should not be used to denote test functions!) It is easy to get a self-consistent system of equations for f together with the pressure gradient, provided we go back, as we did for the relaxed equations, to the trajectorial formulation of the Euler equations,

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

where X_t is volume-preserving in the sense that

$$\int_{\mathcal{A}} \phi(X_t(a))da = \int_{T^d} \phi(x)dx,$$

for all test functions ϕ on $\mathbb{T}^d$. Setting

$$f(t,x,\xi) = \int_{\mathcal{A}} \delta(\xi - \frac{dX_t}{dt}(a))\delta(x - X_t(a))da,$$

we get

$$\partial_t f(t,x,\xi) + \nabla_x \cdot (\xi f(t,x,\xi)) = \nabla_\xi \cdot (\nabla_x p(t,x) f(t,x,\xi)), \quad \int_{\xi\in\mathbb{R}^d} f(t,x,\xi) = 1.$$

Once again, this is an easy consequence of the chain rule, and we only need $\nabla p(t,x)$ to be Lipschitz in $x \in \mathbb{T}^d$ to make it rigorous. Indeed, for every test ϕ function depending only on x, we first find

$$\int_{(x,\xi)\in\mathbb{T}^d\times\mathbb{R}^d} \phi(x) f(t,x,\xi) = \int_{\mathcal{A}} \phi(X_t(a))da = \int_{T^d} \phi(x)dx,$$

and, therefore,

$$\int_{\xi\in\mathbb{R}^d} f(t,x,\xi) = 1.$$

Next, we get for any test function ϕ depending on both x and ξ,

$$\frac{d}{dt}\int_{(x,\xi)\in\mathbb{T}^d\times\mathbb{R}^d} \phi(x,\xi) f(t,x,\xi) = \frac{d}{dt}\int_{\mathcal{A}} \phi(X_t(a), \frac{dX_t}{dt}(a))da$$

$$= \int_{\mathcal{A}} \frac{dX_t}{dt}(a) \cdot (\nabla_x \phi)(X_t(a), \frac{dX_t}{dt}(a))da$$

$$- \int_{\mathcal{A}} (\nabla p)(t, X_t(a)) \cdot (\nabla_\xi \phi)(X_t(a), \frac{dX_t}{dt}(a))da$$

$$= \int_{(x,\xi)\in\mathbb{T}^d\times\mathbb{R}^d} (\xi \cdot \nabla_x \phi(x,\xi) - (\nabla p)(t,x) \cdot \nabla_\xi \phi(x,\xi)) f(t,x,\xi).$$

This "kinetic formulation" of the Euler equations was already introduced in [44] and was, in some sense, the departure points of [43, 46, 48].

2.5 Well-posedness issues

As we have seen, the relaxed Euler equations:

$$\partial_t c(t,x,a) + \nabla_x \cdot (cv(t,x,a)) = 0, \quad \int_a c(t,x,a) = 1,$$

$$(\partial_t(cv) + \nabla_x \cdot (cv \otimes v))(t,x,a) = -c(t,x,a)\nabla_x p(t,x),$$

are very well suited for the "minimizing geodesic problem". It is therefore tempting to think that the relaxed Euler equations, or their kinetic counterpart,

$$\partial_t f(t,x,\xi) + \nabla_x \cdot (\xi f(t,x,\xi)) = \nabla_\xi \cdot (\nabla_x p(t,x) f(t,x,\xi)), \quad \int_{\xi \in \mathbb{R}^d} f(t,x,\xi) = 1,$$

might be good candidates to substitute for the usual Euler equations when we address the initial value problem (IVP), i.e. when we try to get a solution (c, cv) (or f, in kinetic terms), just by prescribing its value at time 0. Unfortunately, it turns out that the relaxed Euler equations are not even well-posed in short time, unless severe restrictions are imposed to the initial conditions $(c_0, c_0 v_0)$ (or f_0 in kinetic terms). Positive and negative results have been obtained in the last 20 years, with many contributors such as Baradat, Bardos and Besse, Brenier, Grenier, Han-Kwan and Iacobelli, Han-Kwan and Rousset, Masmoudi and Wong [17, 19, 49, 155, 157, 158, 191]. Strictly speaking some of these papers, in particular [19, 158], are rather devoted to the “compressible” version of the relaxed Euler equations, which reads, in kinetic terms,

$$\partial_t f(t,x,\xi) + \nabla_x \cdot (\xi f(t,x,\xi)) = \nabla_\xi \cdot (\frac{\nabla_x p}{\rho}(t,x) f(t,x,\xi)), \quad \rho(t,x) = \int_\xi f(t,x,\xi),$$

where the pressure p is a given function of the density ρ.

2.6 Comparison with the Muskat equations

The Euler equations of incompressible inhomogeneous fluids admit a "friction dominated" version which reads (in terms of trajectories)

$$\frac{dX_t}{dt}(a) = -\rho_0(a)G - (\nabla p)(t, X_t(a)), \quad \mathcal{L} \circ X_t^{-1} = \mathcal{L}, \quad \forall t,$$

where we assume, for a moment, that each X_t belongs to $SDiff(D)$. Here, the external force, denoted by G, is a given constant vector in $\mathbb{R}^d$ (typically along the vertical axis, if one considers the gravity force in the simplest possible situation). Notice that the density ρ_0 exclusively features in front of the external force. This corresponds to the so-called "Boussinesq approximation". As a matter of fact, assuming the existence of a velocity field v and a density field ρ such that

$$\frac{dX_t}{dt}(a) = v(t, X_t(a)), \quad \rho(t, X_t(a)) = \rho_0(a),$$

then the equations admit the following "Eulerian" version:

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \nabla \cdot v = 0, \quad v = -\rho G - \nabla p.$$

This set of equations is sometimes called "incompressible porous media equations" or "Muskat's equations" [224], and we will come back to them later in the

course, in subsection 9.1. Notice that they get trivial when there is no external force. (Indeed, in such a case v is both potential and divergence-free.) These equations are very useful for applications (typically, they are the basic equations for "reservoir simulations" in Civil Engineering and Oil Industry). They have been studied in many different ways recently in the mathematical literature, in particular in the framework of convex integration theory. Note that the concept of sub-solutions is not so clearly defined as for the Euler equations, as explained in [224] (that we also quote for the many references it contains).
Anyway, following what we did for the Euler equations, we can easily get a relaxed version for these equations:

Proposition 13. *The Muskat equations admit the following relaxed formulation:*

$$\partial_t c(t,x,a) = \nabla_x \cdot (c(t,x,a)(\rho_0(a)G + \nabla p(t,x)))$$

$$\int_a c(t,x,a) = 1, \quad -\Delta p(t,x) = \nabla_x \cdot \left(\int_a c(t,x,a)\rho_0(a)G \right).$$

Proof (just as before): For all test functions $f = f(x,a)$, we have

$$\frac{d}{dt}\int_{(x,a)} f(x,a)c(t,x,a) = \frac{d}{dt}\int f(X_t(a),a)da$$

$$= \int (\nabla_x f)(X_t(a),a) \cdot \frac{dX_t}{dt}(a)da$$

$$= \int (\nabla_x f)(X_t(a),a) \cdot (-\rho_0(a)G - (\nabla p)(t,X_t(a)))$$

$$= \int_{(x,a)} c(t,x,a)\nabla_x f(x,a) \cdot (-\rho_0(a)G - \nabla p(t,x)).$$

leading to

$$\partial_t c(t,x,a) = \nabla_x \cdot (c(t,x,a)(\rho_0(a)G + \nabla p(t,x))),$$

as announced. Then

$$-\Delta p(t,x) = \nabla_x \cdot \left(\int_a c(t,x,a)\rho_0(a)G \right),$$

immediately follows from $\int_a c(t,x,a) = 1$ by integrating the previous equation with respect to a.
End of proof.

In sharp contrast with the relaxed Euler equations, the relaxed Muskat equations enjoy a well-posedness property for the IVP. This follows from:

Proposition 14. *The relaxed Muskat system admits an extra conservation law for the Boltzmann entropy $\int_a c(t,x,a)$, namely*

$$\partial_t \int_a (c \log c)(t,x,a) + \nabla_x \cdot (\int_a c(t,x,a)\rho_0(a)G) = 0.$$

[This is just a straightforward calculation, since:

$$\partial_t \int_a (c \log c)(t,x,a) = \int_a (1 + \log c(t,x,a)) \nabla_x \cdot ((\rho_0(a)G + \nabla p(t,x))c(t,x,a))$$

$$= -\int_a \nabla_x c(t,x,a) \cdot (\rho_0(a)G + \nabla p(t,x)) = -\nabla_x \cdot (\int_a c(t,x,a)\rho_0(a)G),$$

using that $\int_a c(t,x,a) = 1$.]

Since the Boltzmann entropy is strictly convex in c, the existence of this extra conservation law essentially suffices to guarantee the local well-posedness of the relaxed Muskat equations, following the general theory of entropic system of conservation laws [106] that we will discuss later in the course in section 7.

2.7 Solution of the IVP by convex minimization

It is now quite clear that the relaxed Euler equations are much more adequate for the generalized minimizing geodesic problem, where c is prescribed at the end points $t = 0$ and $t = 1$, for which the solutions are successfully obtained by convex minimization (with a very convincing existence and uniqueness result for the pressure gradient), than for the initial value problem (IVP), when (c, cv) is prescribed at time 0, which is very likely to be ill-posed. Anyway, it seems foolish to solve the IVP problem by a space-time convex minimization technique. Indeed, this may, we are very likely to get optimality equations of space-time elliptic type and therefore ill-posed, although there is a little room left if the convexity is sufficiently degenerate (which is, by the way, the case of the generalized minimizing geodesic problem where the convex functional to be minimized is homogeneous of degree one and, therefore, degenerate). However, as will be discussed at the end of the course, there is a (limited) possibility of that sort which actually involves the cruder concept of sub-solutions we have discussed in the framework of "convex integration" à la De Lellis-Székelyhidi. The idea amounts to minimizing, on a given time interval $[0, T]$,

$$\int_{[0,T]\times\mathbb{T}^d} (\text{trace } M)(dtdx)$$

among all (V, M), where V is square-integrable space-time and M is a bounded Borel space-time measure valued in the set of semi-definite symmetric $d \times d$ matrices, which satisfy $M \geq V \otimes V$ and solve

$$\partial_t V + \nabla \cdot M + \nabla p = 0, \quad \nabla \cdot V = 0,$$

with given initial condition V_0 in the sense

$$\int_Q \partial_t A_i(t,x) V^i(t,x) dtdx + \partial_j A_i(t,x) M^{ij}(dtdx) + \int_D V_0^i(x) A_i(0,x) dx = 0,$$

for all smooth divergence-free vector-fields $A = A(t,x) \in \mathbb{R}^d$ that vanish at $t = T$. It will be shown that:
1) Any smooth solution of the Euler equations can be obtained this way, at least for small enough T.
2) Il may happen that the optimal solution is a classical solution to the Euler equations, but for a $different$ initial condition than V_0! This strange phenomenon is related to the fact that M is just a space-time measure which prevents $V(t,x)$ to be weakly continuous at $t = 0$. Interestingly enough, in some special situations, the resulting solution at time T can be seen as a "relaxed solution", not in the sense we have discussed so far, but rather in the sense developed by Otto [200] for incompressible fluid motions in porous media and recently revisited in [149, 224]. Let us just give an explicit example, with $d = 2$, not on $\mathbb{T}^2$ but rather on $\mathbb{T} \times [-1/2, 1/2]$ (to make the example easier to handle) and we assume $T \leq 1/2$. We take as initial condition

$$V_0(x_1, x_2) = (\mathrm{sign}(x_2), 0),$$

which is an exact, time-independent, discontinuous, trivial solution to the Euler equations, but well known to be "physically unstable" ("Kelvin-Helmholtz instability"). Then, the convex optimization problem provides a completely different solution, which is stationary (i.e. time independent), Lipschitz continuous and explicitly depends on the final time T, namely

$$V_T(x_1, x_2) = \Big(\max(-1, \min(\frac{x_2}{T}, 1)), 0\Big).$$

This looks non sense. However, if we consider this family of stationary solutions as a time dependent solution (the final time T playing the role of the current time), we recover the kind of relaxed solutions advocated by Otto in the (quite different but closely related) framework of incompressible fluid motion in porous media [200, 224]. These topics will be discussed later in the course, in subsection 10.

3 Hidden convexity in some elliptic PDEs

Comme on l'a vu dans le chapitre précédent, le modèle d'Euler des fluides repose essentiellement sur l'EDO

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

où p est la pression qui gouverne la position, notée ici $u(t)$ au temps t, de chaque particule fluide. Cette équation est l'archétype de la mécanique newtonienne (elle décrit la dynamique d'une particule de masse unité, de position $u(t) \in \mathbb{R}^d$ au temps t, dans un potentiel, dépendant du temps, $p(t,x)$ donné sur $\mathbb{R} \times \mathbb{R}^d$). Or cette équation différentielle d'ordre deux a une interprétation variationnelle comme équation d'optimalité d'un problème d'optimisation. C'est aussi le cas d'une classe très large d'équations elliptiques, l'exemple le plus connu étant l'équation de Laplace

$$\Delta u = f$$

d'inconnue u, où f est une fonction donnée sur un domaine $D \subset \mathbb{R}^d$ avec des conditions au bord de D imposées à u (par exemple d'être nulle sur ∂D). Il est bien connu qu'on peut relier cette équation à la minimisation de la fonctionnelle convexe

$$\int_D |\nabla u(x)|^2 + 2f(x)u(x)dx$$

sous des conditions aux limites appropriées. Comme on va le voir dans ce chapitre le principe variationnel peut aller assez loin, jusque, par exemple, à l'équation de Monge-Ampère, version complètement non-linéaire de l'équation de Laplace,

$$\det D^2 u = f.$$

où on note

$$D^2u(x) = \left(\frac{\partial^2 u}{\partial x^i \partial x^j}(t,x), \quad i,j = 1, \cdots, d \right).$$

On pourra en effet relier l'équation de Monge-Ampère, au moins pour une certaine forme de conditions aux limites, à un problème d'optimisation convexe très simple mais passablement caché. Il nous sera assez directement inspirée par l'étude, une fois encore, des équations d'Euler des fluides incompressibles. (Ce qui fait donc de l'équation de Monge-Ampère une nouvelle descendante de celles d'Euler, au côté de celle des ondes ou de la chaleur, comme on l'a vu dans le précédent chapitre).

3.1 Le principe de moindre action pour les EDO

Commençons donc par l'étude variationnelle de l'EDO du second ordre

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

Elle a un caractère "variationnel", au sens qu'elle est la condition d'optimalité de la minimisation d'une certaine "fonctionnelle", à savoir

$$J_{t_0,t_1,p}[u] = \int_{t_0}^{t_1} (\frac{1}{2}|u'(t)|^2 - p(t, u(t)))dt$$

sur des intervalles de temps $[t_0, t_1]$ fixés assez petits (en un sens qu'on va bientôt préciser), parmi toutes les courbes C^1 à valeur dans D à valeurs prescrites en t_0 et t_1. C'est ce qu'on appelle le "principe de moindre action" (qu'Euler attribuait en 1757 à Maupertuis, comme on l'a vu au chapitre précédent).

Theorem 1. *Soit D la clôture d'un ouvert borné convexe lisse de $\mathbb{R}^d$ et soit $[0, T]$ un intervalle de temps borné. Soit p une fonction lisse sur le compact $[0, T] \times D$. Soit une solution u de classe C^2 de l'EDO d'ordre deux*

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

prenant ses valeurs dans D. Alors, pour tout sous-intervalle $[t_0, t_1] \subset [0, T]$ tel que

$$(t_1 - t_0)^2 \sum_{i,j=1}^{d} \partial^2_{x_i x_j} p(t, x) y_i y_j < \pi^2 |y|^2, \quad \forall t \in [t_0, t_1],\ \forall x \in D,\ \forall y \in \mathbb{R}^d \setminus \{0\},$$

on a, pour toute courbe C^1, $t \in [t_0, t_1] \to \tilde{u}(t) \in D$, différente de u,

$$J_{t_0,t_1,p}[\tilde{u}] \ > \ J_{t_0,t_1,p}[u]$$

où J est la fonctionnelle déjà indiquée :

$$J_{t_0,t_1,p}[u] = \int_{t_0}^{t_1} (\frac{1}{2}|u'(t)|^2 - p(t, u(t)))dt.$$

La preuve découle aisément de l'inégalité de Poincaré unidimensionnelle (qu'on peut rapidement prouver par décomposition de Fourier en série de *sinus*).

Lemma 15. *Soit $t_0 < t_1$. Alors pour toute courbe C^1*

$$[t_0, t_1] \to z(t) \in \mathbb{R}^d,$$

telle que $z(t_0) = z(t_1) = 0$

$$\pi^2 \int_{t_0}^{t_1} |z(t)|^2 dt \leq (t_1 - t_0)^2 \int_{t_0}^{t_1} |z'(t)|^2 dt.$$

Preuve du Théorème

Comme p est lisse, il y a une constante $K = K(p) \geq 0$ telle que

$$p(t, \tilde{u}(t)) \leq p(t, u(t)) + \nabla p(t, u(t)) \cdot (\tilde{u}(t) - u(t)) + \frac{1}{2} K(p) |\tilde{u}(t) - u(t)|^2.$$

Comme on suppose D convexe, on peut prendre pour $K(p)$ le sup en (t, x) de la plus grande valeur propre de la matrice symétrique $D^2_x p(t, x)$, i.e.

$$K(p) = \sup_{t,x,|y|=1} \sum_{i,j=1}^{d} \frac{\partial^2 p(t, x)}{\partial x^i \partial x^j} y_i y_j.$$

Par inégalité de Poincaré, on a

$$\int_{t_0}^{t_1} |\tilde{u}(t) - u(t)|^2 dt \leq \frac{(t_1 - t_0)^2}{\pi^2} \int_{t_0}^{t_1} |\tilde{u}'(t) - u'(t)|^2 dt,$$

puisque $\tilde{u}(t_j) = u(t_j)$ pour $j = 0, 1$. Donc

$$\int_{t_0}^{t_1} [p(t, \tilde{u}(t)) - p(t, u(t)) - \nabla p(t, u(t)) \cdot (\tilde{u}(t) - u(t))] dt \leq \int_{t_0}^{t_1} \frac{1}{2} |\tilde{u}'(t) - u'(t)|^2 dt,$$

tant que $t_1 - t_0$ est assez petit de sorte que

$$\frac{(t_1 - t_0)^2}{\pi^2} K(p) \leq 1.$$

Comme u est solution de

$$u''(t) = -(\nabla p)(t, u(t)).$$

on obtient, en intégrant par partie,

$$\int_{t_0}^{t_1} [p(t, \tilde{u}(t)) - p(t, u(t)) - u'(t) \cdot (\tilde{u}'(t) - u'(t))] dt \leq \int_{t_0}^{t_1} \frac{1}{2} |\tilde{u}'(t) - u'(t)|^2 dt,$$

d'où

$$\int_{t_0}^{t_1} [-p(t, u(t)) + \frac{1}{2} |u'(t)|^2] dt \leq \int_{t_0}^{t_1} [-p(t, \tilde{u}(t)) + \frac{1}{2} |\tilde{u}'(t)|^2] dt,$$

ce qui finit la preuve (le cas d'égalité étant laissé en exercice).

3.2 Le principe de moindre action pour les équations d'Euler

On va revenir quelques instants sur les équations d'Euler des fluides incompressibles pour montrer que le principe de moindre action se relève du niveau "individuel" d'une particule fluide accélérée par un potentiel donné p, qu'on vient d'examiner dans la sous-section précédente, au niveau "collectif" du fluide tout entier où la pression n'est plus donnée mais devient elle-même une inconnue du problème. Cela passe par la minimisation de la fonctionnelle

$$\mathbb{J}_{t_0, t_1}[X] = \int_{t_0}^{t_1} \int_D \frac{1}{2} |\partial_t X_t(a)|^2 dx dt$$

où le "flot" $t \to X_t$ est prescrit en $t = t_0$ et $t = t_1$ et contraint de conserver D et son élément de volume à chaque instant t.

Theorem 2. *Soit une solution (X, p) des équations d'Euler au sens de*

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

$$\int_D \phi(t, X_t(a)) dx = \int_D \phi(x) dx, \quad \forall \phi \in C^0(\mathbb{R}^d), \quad \forall t.$$

Soit un intervalle de temps $[t_0, t_1]$ suffisamment petit de sorte que

$$\frac{(t_1 - t_0)^2}{\pi^2} K(p) \leq 1$$

où $K(p)$ est le sup en (t, x) de la plus grande valeur propre de la matrice symétrique $D_x^2 p(t, x)$.

Alors, pour toute application $(t, a) \in [t_0, t_1] \times D \to \tilde{X}_t(a) \in \mathbb{R}^d$ *de carré sommable, conservant* D *et son élément de volume au sens*

$$\int_D \phi(\tilde{X}_t(a))da = \int_D \phi(x)dx, \quad \forall \phi \in C^0(\mathbb{R}^d), \quad \forall t \in [t_0, t_1],$$

telle que

$$\tilde{X}_{t_0} = X_{t_0}, \quad \tilde{X}_{t_1} = X_{t_1},$$

différente de X*, on a*

$$\mathbb{J}_{t_0,t_1}[\tilde{X}] > \mathbb{J}_{t_0,t_1}[X]$$

où

$$\mathbb{J}_{t_0,t_1}[\tilde{X}] = \int_{t_0}^{t_1} \int_D \frac{1}{2}|\partial_t \tilde{X}_t(a)|^2 dadt.$$

Remarque

Ceci nous permet, en passant (car ce n'est pas le point central de ce chapitre, principalement consacré aux équations elliptiques), de donner une justification, en termes élémentaires (sans connaissances requises en géométrie riemannienne, dans le simple cas d'un domaine convexe borné de $\mathbb{R}^d$) de la fameuse interprétation géométrique des équations d'Euler des fluides incompressibles, par V.I. Arnold en 1966 [selon laquelle, dans le cas d'une variété riemannienne compacte D, elles décrivent les géodésiques -à vitesse constante- du groupe de Lie (formel) de dimension infinie des difféomorphismes conservant le volume de la variété, pour la métrique L^2 sur son algèbre de Lie (formelle), constituée des champs de vecteurs à divergence nulle].

Preuve du Théorème

La démonstration découle immédiatement du Théorème 1. En effet, pour chaque $a \in D$ fixé, on a, en posant $u(t) = X_t(a)$ et $\tilde{u}(t) = \tilde{X}_t(a)$,

$$\int_{t_0}^{t_1} [-p(t, u(t)) + \frac{1}{2}|u'(t)|^2]dt \leq \int_{t_0}^{t_1} [-p(t, \tilde{u}(t)) + \frac{1}{2}|\tilde{u}'(t)|^2]dt,$$

et donc

$$\int_{t_0}^{t_1} [-p(t, X_t(a))) + \frac{1}{2}|\partial_t X_t(a)|^2]dt \leq \int_{t_0}^{t_1} [-p(t, \tilde{X}_t(a))) + \frac{1}{2}|\partial_t \tilde{X}_t(a)|^2]dt.$$

En intégrant en a et en prenant en compte la conservation des volumes par à la fois X et $\tilde{X}$, on trouve l'inégalité (large) voulue :

$$\int_D \int_{t_0}^{t_1} \frac{1}{2}|\partial_t X_t(a)|^2 dtda \leq \int_D \int_{t_0}^{t_1} \frac{1}{2}|\partial_t \tilde{X}_t(a)|^2 dtda$$

(et on laisse le cas d'égalité en exercice).
Fin de la preuve.

On peut aller encore un peu plus loin, et ce sera utile pour l'étude de l'équation de Monge-Ampère, en montrant que, étant données les valeurs de X_t en $t = t_0$ et $t = t_1$, on peut retrouver le champ de pression p comme maximiseur d'une certaine fonctionnelle convexe.

Theorem 3. *Soit un intervalle de temps* $[t_0, t_1]$ *suffisamment petit de sorte que*

$$\frac{(t_1 - t_0)^2}{\pi^2} K(p) \leq 1$$

où $K(p)$ *est le sup en* (t, x) *de la plus grande valeur propre de la matrice symétrique* $D_x^2 p(t, x)$. *Alors, le champ de pression* p *maximise la fonctionnelle*

$$\mathbb{K}_{t_0,t_1}[p] = \int_{t_0}^{t_1} \int_D p(t, x) dx dt + \int_D K_{t_0,t_1,p}(X_{t_0}(a), X_{t_1}(a)) da,$$

où

$$K_{t_0,t_1,p}(u_0, u_1) = \inf\{\int_{t_0}^{t_1} (\frac{1}{2}|u'(t)|^2 - p(t, u(t)))dt,\ u \in C^1,\ u(t_0) = u_0,\ u(t_1) = u_1\}$$

Preuve du Théorème

Soit $\tilde{p}$ un "compétiteur" pour p. Par définition de

$$K_{t_0,t_1,\tilde{p}}(u_0, u_1) = \inf\{\int_{t_0}^{t_1} (\frac{1}{2}|u'(t)|^2 - \tilde{p}(t, u(t)))dt,\ u \in C^1,\ u(t_0) = u_0,\ u(t_1) = u_1\},$$

on a, pour tout $a \in D$ fixé,

$$K_{t_0,t_1,\tilde{p}}\left(X_{t_0}(a), X_{t_1}(a)\right) \leq \int_{t_0}^{t_1} \left(\frac{1}{2}|\partial_t X_t(a)|^2 - \tilde{p}(t, X_t(a))\right) dt.$$

En intégrant en $a \in D$ cette relation on obtient aussitôt

$$\begin{aligned}\int_D K_{t_0,t_1,\tilde{p}}\left(X_{t_0}(a), X_{t_1}(a)\right) da &\leq \int_D \int_{t_0}^{t_1} \left(\frac{1}{2}|\partial_t X_t(a)|^2 - \tilde{p}(t, X_t(a))\right) dt da \\ &= \int_D \int_{t_0}^{t_1} \left(\frac{1}{2}|\partial_t X_t(a)|^2 - \tilde{p}(t, a)\right) dt da\end{aligned}$$

(en utilisant que X_t conserve les volumes). Pour p on obtient, pour chaque $a \in D$, l'égalité (et non plus l'inégalité)

$$K_{t_0,t_1,p}\left(X_{t_0}(a), X_{t_1}(a)\right) = \int_{t_0}^{t_1} \left(\frac{1}{2}|\partial_t X_t(a)|^2 - p(t, X_t(a))\right) dt$$

(grâce au Théorème 1, en utilisant la petitesse de l'intervalle $[t_0, t_1]$) et donc, en intégrant en a,

$$\int_D K_{t_0,t_1,p}\left(X_{t_0}(a), X_{t_1}(a)\right) da = \int_D \int_{t_0}^{t_1} \left(\frac{1}{2}|\partial_t X_t(a)|^2 - p(t, a)\right) dt da.$$

En soustrayant l'égalité portant sur p de l'inégalité portant sur $\tilde{p}$, on obtient

$$\int_{t_0}^{t_1} \int_D \tilde{p}(t,x)dxdt + \int_D K_{t_0,t_1,\tilde{p}}(X_{t_0}(a), X_{t_1}(a))da$$

$$\leq \int_{t_0}^{t_1} \int_D p(t,x)dxdt + \int_D K_{t_0,t_1,p}(X_{t_0}(a), X_{t_1}(a))da,$$

i.e. le résultat voulu, à savoir :

$$\mathbb{K}_{t_0,t_1}[\tilde{p}] \leq \mathbb{K}_{t_0,t_1}[p].$$

Ainsi on a montré que le champ de pression p satisfait un principe de maximisation qu'on peut considérer comme "dual" du principe de minimisation que satisfait X au travers du Théorème 2. L'avantage de ce nouveau problème est qu'il s'agit, étant donnés un intervalle $[t_0, t_1]$ et deux transformations X_{t_0} et X_{t_1} conservant D et son volume, d'un problème de maximisation *concave* en p alors que pour X le problème correspondant de minimisation n'est pas convexe à cause de la contrainte de conservation des volumes qui n'est pas du tout convexe ! De plus, on observe que le problème en p n'implique aucune dérivée partielle contrairement au problème en X où figure ∂_t. Nous reviendrons sur La résolution de ce problème dans le prochain chapitre.

3.3 Equation de Monge-Ampère réelle et transport optimal

Le problème de maximisation concave (permettant de retrouver la pression dans le cadre des équations d'Euler) que nous venons de voir nous inspire la considération d'un problème voisin mais beaucoup plus simple (suivant le principe "qui peut le plus peut le moins"), à savoir la maximisation de la fonctionnelle

$$\phi \to \int_{\mathbb{R}^d} \phi(x)\rho_0(x)dx + \int_{\mathbb{R}^d} \inf_{x\in\mathbb{R}^d} \left(\frac{1}{2}|y-x|^2 - \phi(x)\right) \rho_1(y)dy,$$

où les fonctions ρ_0 et ρ_1, positives, à support compact et d'intégrale de Lebesgue égale à 1, sont données sur $\mathbb{R}^d$. Ce problème va nous permettre de résoudre l'équation elliptique très fameuse en géométrie (riemannienne et kählerienne) et "complètement non-linéaire", de Monge-Ampère (réelle)

$$\rho_1(x + \nabla\phi(x))\det(\mathbb{I}_d + D^2\phi(x)) = \rho_0(x)$$

(qui a permis à Minkowski, il y a un peu plus d'un siècle, de montrer qu'on pouvait reconstruire des hypersurfaces convexes par la seule connaissance de leur courbure gaussienne). On va utiliser pour cela une technique assez élémentaire d'analyse convexe, dérivée de la théorie de Monge-Kantorovich du "transport optimal" -dans laquelle se sont illustrés deux médailles Fields récentes (Villani 2010 et Figalli 2018). On observera encore que la fonctionnelle à maximiser est bien concave et ne comporte aucune dérivée partielle, alors que l'EDP en jeu est particulièrement non-linéaire avec des non-linéarités portant (au moins quand $d > 1$) sur les dérivées partielles de plus haut degré. On parle dans ce cas d'EDP totalement (ou complètement) non-linéaire ("fully nonlinear" en anglais).

Theorem 4. *Soit B une boule fermée de $\mathbb{R}^d$ centrée en 0. Soient deux mesures de probabilité boréliennes sur B, μ_0 et μ_1. On suppose que μ_0 est a.c. par rapport à la mesure de Lebesgue, i.e. $\mu_0(dx) = \rho_0(x)dx$ avec $\rho_0 \geq 0$ dans $L^1(B)$ d'intégrale 1. Alors, il existe une unique application borélienne T envoyant μ_0 sur μ_1 qui s'écrive $T(x) = \nabla a(x)$, $\rho_0(x)dx$ presque partout, pour une fonction convexe Lipschitzienne a sur B.*

Remarque

On peut dire qu'on ainsi résolu, en un sens généralisé, dans le cas où $\mu_1(dy) = \rho_1(y)dy$ avec $\rho_1 \in L^1(B)$, l'équation de Monge-Ampère (réelle)

$$\rho_1(\nabla a(x))\det(D^2 a(x)) = \rho_0(x),$$

avec a convexe Lipschitzienne sur B. En effet, en supposant que le changement de variable $x \in B \to y = \nabla a(x) \in B$, $dy = \det(D^2 a(x))dx$ est valide, on a, pour tout $u \in C^0(B)$,

$$\int_B u(y)\rho_1(y)dy = \int_B u(\nabla a(x))\rho_1(\nabla a(x))\det(D^2 a(x))dx = \int_B u(\nabla a(x))\rho_0(x)dx,$$

ce qui dit bien que $x \to \nabla a(x)$ envoie $\rho_0(x)dx$ sur $\rho_1(y)dy$ quand l'équation de MA est satisfaite.
Comme on va le voir, la démonstration du théorème 4 passe par la résolution du problème d'optimisation (dit de Monge-Kantorovich)

$$\inf\{\int_B a(x)\mu_0(dx) + \int_B b(y)\mu_1(dy), \quad (a,b) \in C^0(B) \times C^0(B)\},$$

sous contrainte $a(x) + b(y) \geq x \cdot y$, $\forall x, y \in B$. Ainsi on va résoudre une EDP totalement non-linéaire à l'aide d'un "programme linéaire" où ne figure aucune dérivée partielle !

Preuve du Théorème 4

La preuve n'utilise que deux théorèmes assez classiques d'analyse : celui de dualité convexe de Fenchel-Rockafellar (qu'on trouve dans le livre d'analyse fonctionnelle de Brezis, au premier chapitre, comme corollaire du théorème de Hahn-Banach) et celui de Rademacher montrant la différentiabilité presque partout des fonctions Lipschitz continues.

On introduit l'espace de Banach

$$\mathcal{W} = C^0(B \times B)$$

équipé de la norme du sup et on se donne une fonction continue c sur $B \times B$. (Plus tard, on particularisera $c(x, y) = x \cdot y$.) On introduit aussi les fonctions convexes sur $\mathcal{W}$, à valeurs dans $]-\infty, +\infty]$, respectivement définies pour chaque $w \in \mathcal{W}$ par :

$$\Phi(w) = 0, \text{ si } w \geq c, \quad +\infty \text{ sinon},$$

$$\Psi(w) = \int_B a(x)\mu_0(dx) + \int_B b(y)\mu_1(dy) \text{ si } w = a \oplus b,$$

pour des fonctions continues a et b sur B, et $+\infty$ autrement. [Notons que Ψ est défini sans ambigüité car μ_0 et μ_1 ont même masse.] Observons qu'il y a au moins un point $w \in \mathcal{W}$ où Φ est continue et Ψ bornée. [Prendre par exemple pour w la fonction constante égale à 1 plus le sup de c sur $B \times B$.] Comme Φ et Ψ sont convexes, cela suffit pour appliquer le théorème de Fenchel-Rockafellar (comme énoncé dans le livre de Brezis d'anayse fonctionnelle, par exemple). Il nous dit que

$$\inf\{\Phi(w) + \Psi(w),\ w \in \mathcal{W}\} = \max\{-\Phi^*(-\mu) - \Psi^*(\mu),\ \mu \in \mathcal{W}'\}$$

où $\mathcal{W}'$ est le Banach dual $\mathcal{W}$, i.e. celui des mesures de Borel μ bornées sur $B \times B$ (par le théorème de Riesz), et où Φ^*, Ψ^* sont les transformées de Legendre-Fenchel de Φ et Ψ, i.e.

$$\Phi^*(\mu) = \sup\{< \mu, w > -\Phi(w), \quad w \in \mathcal{W}\}$$

$$\Psi^*(\mu) = \sup\{< \mu, w > -\Psi(w), \quad w \in \mathcal{W}\},$$

où le "crochet de dualité" est ici défini par :

$$< \mu, w >= \int_{B \times B} w(x, y)\mu(dx, dy), \quad \forall w \in \mathcal{W}, \quad \forall \mu \in \mathcal{W}'.$$

La notation "max" est utilisée à dessein pour indiquer que le sup est atteint dans le second membre (alors que ce n'est pas forcément le cas de l'inf du côté gauche).

Remarque "culturelle" : Le lecteur "averti" pourra reconnaitre dans le théorème de Fenchel-Rockafellar l'équivalent $(\max, +)$ de celui de Plancherel en analyse de Fourier, la transformation de Legendre-Fenchel se substituant à celle de Fourier par la même occasion. Notons que le passage de $(+, \times)$ à $(\max, +)$ correspond aussi à celui de la géométrie algébrique à la géométrie "tropicale" !

Calculons maintenant Φ^* et Ψ^*. On obtient d'abord $\Phi^*(-\mu) = +\infty$, sauf si $\mu \geq 0$, auquel cas

$$\Phi^*(-\mu) = -\int_{B \times B} c(x, y)\mu(dx, dy).$$

Ensuite $\Psi^*(\mu) = +\infty$, sauf si les deux projections de μ sont respectivement μ_0 et μ_1, auquel cas $\Psi^*(\mu) = 0$. Ainsi on a obtenu qu'il existe $\mu_{\text{opt}} \geq 0$, de projections

μ_0 et μ_1 qui maximise $\int_{B\times B} c(x,y)\mu(dx,dy)$ parmi toutes les mesures de Borel positives sur $B\times B$ dont les deux projections sont respectivement μ_0 et μ_1. On a aussi la relation de dualité :

$$\int_{B\times B} c(x,y)\mu_{\mathrm{opt}}(dx,dy) = \inf\{\int_B a(x)\mu_0(dx) + \int_B b(y)\mu_1(dy), \quad a\oplus b \geq c\}.$$

A priori cet inf n'est pas atteint. Considérons pourtant une suite minimisante (a_n, b_n). Assez miraculeusement, on peut construire une nouvelle suite minimisante $(\tilde{a}_n, \tilde{b}_n)$. encore plus performante, juste en posant

$$\tilde{b}_n(y) = \sup_{x\in B} c(x,y) - a_n(x),$$

$$\tilde{a}_n(x) = \sup_{y\in B} c(x,y) - \tilde{b}_n(y).$$

(En effet on a $\tilde{b}_n \leq b_n$, $\tilde{a}_n \leq a_n$ et $\tilde{a}_n \oplus \tilde{b}_n \geq c$.) Cette nouvelle suite est uniformément équicontinue sur B (sachant que c est continue sur $B\times B$ qui est compact). On la note de nouveau (a_n, b_n) pour alléger l'écriture. Comme on peut ajouter une constante arbitraire à a_n et retrancher la même de b_n, on peut supposer que les $\tilde{a}_n$ et les $\tilde{b}_n$ sont uniformément bornées sur B. (En effet, on peut ajuster a_n se sorte que le sup de $x\to c(x,0)-a_n(x)$ sur B est égal à 0, ce qui assure que $a_n \geq \inf c$ et $\tilde{b}_n(0)=0$. Il s'ensuit que les $|\tilde{b}_n|$ sont uniformément bornées par une constante R, puisqu'uniformément équicontinues. Par définition, les $\tilde{a}_n$ sont bornées supérieurement par $R+\sup c$ et inférieurement par $\inf c$.) On peut donc appliquer le théorème d'Ascoli (sous sa forme la plus simple) pour s'assurer que (a_n, b_n) converge uniformément vers une limite (a,b) sur B. On peut même s'assurer que

$$a(x) = \sup_{y\in B} c(x,y) - b(y)$$

(suivant le procédé déjà utilisé). On voit tout de suite que (a,b) minimise la fonctionnelle (qui est clairement continue sur $C(B)\times C(B)$)

$$(a,b) \to \int_B a(x)\mu_0(dx) + \int_B b(y)\mu_1(dy)$$

parmi les couples (a,b) tels que $a\oplus b\geq c$. On a donc

$$\int_{B\times B} c(x,y)\mu_{\mathrm{opt}}(dx,dy) = \int_B a(x)\mu_0(dx) + \int_B b(y)\mu_1(dy),$$

d'où l'on déduit

$$\int_{B\times B} (a(x)+b(y)-c(x,y))\mu_{\mathrm{opt}}(dx,dy) = 0,$$

puisque μ_0 et μ_1 sont les deux projections de μ_{opt}. Comme μ_{opt} est une mesure positive, cela impose

$$a(x)+b(y) = c(x,y)$$

pour $\mu_{\text{opt}}-$presque tous x, y dans B.

Supposons dorénavant que $c(x,y) = x \cdot y$ et que μ_0 est absolument continue par rapport à la mesure de Lebesgue et s'écrit donc

$$\mu_0(dx) = \rho_0(x)dx,$$

pour une fonction Lebesgue intégrable $\rho_0 \geq 0$ sur B, d'intégrale 1. On a ainsi

$$a(x) = \sup_{y \in B} \quad x \cdot y - b(y)$$

ce qui montre que a est à la fois Lipschitz et convexe sur B. Le théorème de Rademacher nous dit qu'une fonction Lipschitzienne sur B est différentiable Lebesgue presque partout à l'intérieur de B. Sachant que B est lisse, sa frontière est de mesure de Lebesgue nulle. On voit donc que l'ensemble des points x de B qui sont soit au bord de B soit à l'intérieur de B et où a n'est pas différentiable est de mesure de Lebesque nulle et donc, puisque μ_0 est supposée a.c. par rapport à Lebesgue, aussi de μ_0-mesure nulle. Comme μ_{opt} admet μ_0 pour projection, on déduit que, pour μ_{opt} presque tout couple $(x^*, y^*) \in B \times B$, x^* est à l'intérieur de B et a est différentiable en x^*. On peut en plus supposer que

$$a(x^*) + b(y^*) = x^* \cdot y^*,$$

puisque, comme on l'a vu, cette propriété est vraie μ_{opt} presque partout. Comme on a

$$a(x) + b(y^*) \geq x \cdot y^*$$

pour tout $x \in B$, on en tire que x^* est un point de minimum de la fonction $x \to a(x) - x \cdot y^*$. Donc, en différentiant, on obtient que

$$\nabla a(x^*) = y^*.$$

Cette propriété est donc vraie μ_{opt} presque partout ce qui implique

$$\mu_{\text{opt}}(dx, dy) = \delta(y - \nabla a(x))\rho_0(x)dx,$$

au sens précis que

$$\int_{B \times B} w(x,y)\mu_{\text{opt}}(dx, dy) = \int_B w(x, \nabla a(x))\rho_0(x)dx, \quad \forall w \in C(B \times B).$$

(Notons, en passant, que cela implique l'unicité de la solution optimale μ_{opt}.) Par projection (i.e. en prenant $w(x,y) = u(y)$) , on en déduit

$$\int_B u(y)\mu_1(dy) = \int_B u(\nabla a(x))\rho_0(x)dx, \quad \forall u \in C(B),$$

ce qui nous dit exactement que $x \to \nabla a(x)$ envoie la mesure $\rho_0(x)dx$ sur la mesure $\mu_1(dy)$. Comme a est Lipschitz et convexe, on a montré la partie d'existence du théorème 4.

Reste à montrer l'unicité. Pour cela, soit $\tilde{a}$ une fonction convexe Lipschitzienne telle que $x \to y = \nabla\tilde{a}(x)$ envoie $\rho_0(dx)$ sur $\mu_1(dy)$ et posons

$$\tilde{b}(y) = \sup_{x\in B} \quad x \cdot y - \tilde{a}(x), \quad y \in B.$$

Montrons d'abord que

$$\tilde{a}(x) + \tilde{b}(\nabla\tilde{a}(x)) = x \cdot \nabla\tilde{a}(x)$$

est vraie pour Lebesgue presque tout $x \in B$. [En effet, notons d'abord que presque tout $x^* \in B$ est un point intérieur de B et de différentiabilité pour $\tilde{a}$ (par Rademacher). Posons $y^* = \nabla\tilde{a}(x^*)$. Observons que y^* est à valeur dans B, puisque $\nabla\tilde{a}$ envoie $\rho(x)dx$ sur $\mu_1(dy)$ et que ses deux mesures ont leur support dans B (qui est fermé). La fonction $x \in B \to x \cdot y^* - \tilde{a}(x)$ est Lipschitz concave sur B, dérivable en $x = x^*$ qui est à l'intérieur de B, et de dérivée nulle. Elle y atteint donc son maximum, qui est, par définition $\tilde{b}(y^*)$. On a donc $\tilde{b}(y^*) = x^* \cdot y^* - \tilde{a}(x^*)$. Comme $y^* = \nabla\tilde{a}(x^*)$, on a obtenu le résultat voulu.]
Posons maintenant

$$\mu(dx, dy) = \delta(y - \nabla\tilde{a}(x))\rho_0(x)dx.$$

On a donc

$$\int_{B\times B} x \cdot y\mu(dx, dy) = \int_B x \cdot \nabla\tilde{a}(x)\rho_0(x)dx$$

$$= \int_B (\tilde{a}(x) + \tilde{b}(\nabla\tilde{a}(x))\rho_0(x)dx = \int_{B\times B} (\tilde{a}(x) + \tilde{b}(y))\mu(dx, dy)$$

$$= \int_{B\times B} (\tilde{a}(x) + \tilde{b}(y))\mu_{\text{opt}}(dx, dy)$$

(puisque μ_{opt} et μ ont les mêmes projections)

$$\geq \int_{B\times B} x \cdot y\mu_{\text{opt}}(dx, dy)$$

(puisque $\tilde{a}(x) + \tilde{b}(y) \geq x \cdot y$). Donc μ est aussi optimale que μ_{opt} et, comme on l'a vu, cette dernière est unique solution optimale. On a donc, par définition de μ :

$$\delta(y - \nabla\tilde{a}(x))\rho_0(x)dx = \mu(dx, dy) = \mu_{\text{opt}}(dx, dy) = \delta(y - \nabla a(x))\rho_0(x)dx,$$

et ceci n'est évidemment possible que si $\nabla\tilde{a}(x) = \nabla a(x)$ pour $\rho_0(x)dx$ presque tout x. On a ainsi obtenu l'unicité dans le théorème 4.

3.4 Application au problème de la “meilleure constante de Sobolev”

On va maintenant s'intéresser au problème de minimisation non convexe

$$I(U,p,q) = \inf\{\int_U |\nabla u(x)|^p dx, \quad u \in C_c^\infty(U), \ \text{t.q.} \ \int_U |u(x)|^q dx = 1\ \}$$

où $p, q \in]1, +\infty[$ et U est un ouvert de $\mathbb{R}^d$ et on va voir comment lui appliquer le théorème de transport optimal qu'on vient de prouver dans le sous-section préćédente (et qui repose sur un problème de maximisation concave) pour le résoudre.
On voit d'abod assez facilement, en effectuant des changements de variable linéaires du type $x \to rx + a$ avec $r > 0$ et $a \in \mathbb{R}^d$ sur les fonctions $u \in C_c^\infty(U)$, que :
i) dans le cas $U = \mathbb{R}^d$, $I(U,p,q) = 0$ sauf si $1 - d/p = 0 - d/q$;
ii) dans le cas où U est borné (et où on utilise uniquement des rétractions avec $r > 1$) $I(U,p,q) = 0$ sauf si

$$1 - d/p \geq 0 - d/q.$$

(Retenons l'indice $1 - d/p$ qui signifie "une dérivée dans L^p en d dimensions d'espace". Il permet de comparer les espaces de Sobolev entre eux.) Dans le cas U borné avec exposant "sous-critique", i.e. $1 - d/p > 0 - d/q$, on peut assez facilement généraliser la traditionnelle méthode de compacité en utilisant qu'une suite u_n telle que

$$\sup_n \int_U |\nabla u_n(x)|^p dx < +\infty$$

admet une sous-suite fortement convergente dans $L^q(U)$. On déduit au moins l'existence d'une solution généralisée optimale, cela dans l'espace de Sobolev obtenu par complétion de $C_c^\infty(D)$ pour la norme

$$u \to ||u||_{L^q(U)} + ||\nabla u||_{L^p(U)}.$$

Il est aussi assez facile de voir qu'une telle solution vérifie forcément au sens des distributions dans U l'équation :

$$-\nabla(|\nabla u|^{p-2}\nabla u) = \lambda u|u|^{q-2}$$

où la constante λ (souvent appelée "valeur propre non-linéaire") est à ajuster pour assurer la normalisation $||u||_{L^q(U)} = 1$. En particulier, dans le cas le plus utile $p = 2$, on trouve l'équation semi-linéaire

$$-\Delta u = \lambda u|u|^{q-2}.$$

Dans le cas "critique" $1-d/p = 0-d/q$, il est encore facile de voir que $I(U,p,q)$ ne dépend pas de U! Plus subtilement (et cela sort du cadre du cours en faisant appel à la théorie de la “concentration-compacité” [175]), lorsque U est borné, il n'y a pas de solution optimale, même dans l'espace complété. On peut même montrer que les suites minimisantes u_n ont la curieuse propriété, à extraction de sous-suite près, de se concentrer au sens qu'il existe un point x_∞ de U telle que $|u_n|^q$ converge au sens des mesures de Borel vers la masse de Dirac au point x_∞. C'est le fameux phénomène de "concentration", ou encore de "formation de bulles", si fréquent en analyse géométrique (y compris dans la résolution de la conjecture de Poincaré par Perelman). Pour un résultat positif, on va donc se concentrer sur le seul cas $U = \mathbb{R}^d$ et montrer

Theorem 5. *Dans le cas critique* $1 - d/p = 0 - d/q$,

$$I(\mathbb{R}^d, p, q) = \inf\{\int_{\mathbb{R}^d} |\nabla u(x)|^p dx, \quad u \in C_c^\infty(\mathbb{R}^d), \ \text{t.q.} \ \int_{\mathbb{R}^d} |u(x)|^q dx = 1 \}$$

est atteint, par une unique solution u *(modulo translations et dilatations), dans le Banach* E*, complété de* $C_c^\infty(\mathbb{R}^d)$ *pour la norme*

$$||u||_E = ||u||_{L^q(\mathbb{R}^d)} + ||\nabla u||_{L^p(\mathbb{R}^d)}.$$

Du même coup, l'équation

$$-\nabla(|\nabla u|^{p-2}\nabla u) = \lambda u|u|^{q-2}$$

est uniquement (modulo translations et dilatations) résolue sur $\mathbb{R}^d$ *dans l'espace* E*, la constante* λ *étant à ajuster de sorte que*

$$||u||_{L^q(\mathbb{R}^d)} = 1.$$

Il y a plusieurs preuves possibles, en particulier par "concentration-compacité" [175]. Une preuve étonnamment simple (à condition de négliger pas mal d'aspects techniques) découle directement du théorème 4 sur la résolution de l'équation de Monge-Ampère. Elle a été trouvée par Cordero-Erausquin, Nazaret et Villani [101]. L'idée de base est de considérer, pour deux fonctions u et v dans $C_c^\infty(\mathbb{R}^d)$ telles que $||u||_{L^q(\mathbb{R}^d)} = ||v||_{L^q(\mathbb{R}^d)} = 1$, les deux mesures boréliennes de probabilités

$$F(x)dx = |u(x)|^q dx, \quad G(y)dy = |v(y)|^q dy.$$

Comme on l'a vu plus haut, on peut construire une unique application borélienne T qui transporte la première mesure vers la seconde et s'écrive, pour $F(x)dx$ presque tout x,

$$T(x) = \nabla\Phi(x),$$

où Φ est convexe Lipschitzienne sur $\mathbb{R}^d$. De plus, en un sens généralisé, Φ est solution de l'équation de Monge-Ampère

$$G(\nabla\phi(x))\det(D^2\Phi(x)) = F(x).$$

On va simplement évaluer l'intégrale

$$J = \int_{\mathbb{R}^d} G(y)^{1-1/d} dy$$

et, assez magiquement, tout va découler, (existence, unicité et formule explicite pour les solutions du problème de la meilleure constante de Sobolev) quasi-automatiquement de deux inégalités très élémentaires. Celle de Young

$$\frac{|a|^p}{p} + \frac{|b|^{p'}}{p'} \geq a \cdot b, \ \ \forall a, b \in \mathbb{R}^d, \quad 1/p' + 1/p = 1, \quad p \in]1, \infty[$$

(avec égalité si et seulement si $b = a|a|^{p-2}$ ou $a = b|b|^{p'-2}$) et celle qui dit que, pour toute suite finie de réels positifs, la moyenne géométrique est toujours inférieure à la moyenne arithmétique avec égalité si et seulement si les dits réels sont tous égaux.

Par construction de $T = \nabla\Phi$, on a d'abord

$$J = \int_{\mathbb{R}^d} G(y)^{1-1/d} dy = \int_{\mathbb{R}^d} G(\nabla\Phi(x))^{-1/d} F(x) dx$$
$$= \int_{\mathbb{R}^d} \det(D^2\Phi(x))^{1/d} F(x)^{1-1/d} dx.$$

(Ici la preuve mériterait d'être précisée car l'équation de M.-A. n'est pas a priori satisfaite au sens classique. On passera ici sur ce point passablement technique qui utilise la notion de solution au sens d'Alexandrov.) C'est maintenant qu'intervient l'inégalité de la moyenne géométrique, sachant que la convexité de Φ assure (au moins formellement) que les valeurs propres de $D^2\Phi$ sont réelles et positives (au sens large), ce qui nous donne, ponctuellement,

$$\det(D^2\Phi(x))^{1/d} \leq 1/d \; \Delta\Phi(x).$$

On en tire $J \leq \tilde{J}$ où

$$\tilde{J} = 1/d \int_{\mathbb{R}^d} \Delta\Phi(x) F(x)^{1-1/d} dx$$

$$= -1/d \int_{\mathbb{R}^d} \nabla\Phi(x) \cdot \nabla(F(x)^{1-1/d}) dx$$

(par intégration par partie)

$$= -s/d \int_{\mathbb{R}^d} \nabla\Phi(x) \cdot u(x)|u(x)|^{s-2} \nabla u(x) dx$$

(en posant $s = (1 - 1/d)q$ et par définition de $F = |u|^q$)

$$\leq s/d ||\nabla u||_{L^p(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} |u(x)|^{(s-1)p'} |\nabla\Phi(x)|^{p'} dx \right)^{1/p'}$$

(par Young-Hölder, avec $1/p' = 1 - 1/p$)

$$= s/d ||\nabla u||_{L^p(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} F(x) |\nabla\Phi(x)|^{p'} dx \right)^{1/p'},$$

(notant que $(s-1)p' = q$ et sachant que $F = |u|^q$)

$$= s/d||\nabla u||_{L^p(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} G(y)|y|^{p'} dy \right)^{1/p'},$$

(puisque $G(y)dy$ est l'image par $T = \nabla\Phi$ de $F(x)dx$). Ainsi, on a obtenu que, pour tous u et v de norme 1 dans L^q,

$$\int_{\mathbb{R}^d} |v(y)|^s dy \leq s/d||\nabla u||_{L^p(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} |v(y)|^q |y|^{p'} \right)^{1/p'}$$

avec $s = (1-1/d)q$ et ceci s'étend à u et v pris dans le Banach complété E. Notons de plus que cette inégalité devient égalité si et seulement si l'inégalité géométrique et celle de Hölder sont toutes deux saturées. On trouve alors (à la suite d'un calcul un peu laborieux) une constante $r > 0$ et un point x_0 tels que $T(x) = (x - x_0)r$, $u(x) = r^{-d} v((x - x_0)r)$ et, enfin, $u(x) = (\mu + |x - x_0|^\alpha)^\beta \nu$ pour des constantes α, β, μ, ν à ajuster en fonction de p et d via q et s (en fait $\alpha = p'$ et $\beta = 1 - d/p = -d/q$). [Observons que, concernant u et v, on est bien sorti de l'espace $C_c^\infty(\mathbb{R}^d)$ pour se retrouver dans le complété E.]

Ceci se traduit donc par la remarquable relation suivante, de dualité entre deux problèmes d'optimisation *non convexes* (pas question, a priori, d'appliquer ici le théorème de dualité convexe de Fenchel-Rockafellar!),

$$\max_{v \in S_1(L^q)} \frac{\int_{\mathbb{R}^d} |v(y)|^s dy}{\left(\int_{\mathbb{R}^d} |v(y)|^q |y|^{p'} dy \right)^{1/p'}} = s/d \min_{u \in S_1(L^q)} ||\nabla u||_{L^p(\mathbb{R}^d)},$$

$$s = (1-1/d)q, \ 1 - d/p = -d/q,$$

où $S_1(L^q)$ note la sphère unité de L^q intersectée avec E. L'existence, l'unicité (modulo translations et dilatations) de solutions (dans le Banach complété E) au problème de la meilleure constante de Sobolev, et leur forme explicite, ne sont en fait que des corollaires de cette relation de dualité non-convexe.

4 Analysis of the optimal incompressible transport problem

This chapter is entirely devoted to the analysis of the relaxed minimizing geodesic problem, already presented in subsection 2.2 of chapter 2, that we can call, as well, the "optimal incompressible transport". Let us recall pour un couple quelconque de mesures $(c, q) \in \left(C^0(D^2;\ \mathbb{R} \times \mathbb{R}^d)\right)'$

$$K(c,q) = \sup\{\int_{D^2} A(x,a)dc(x,a) + B(x,a) \cdot dq(x,a);$$

$$(A,B) \in C^0(D^2;\ \mathbb{R} \times \mathbb{R}^d) \text{ s.t. } 2A + |B|^2 \leq 0\},$$

qui a l'avantage de définir une fonction convexe s.c.i à valeurs dans $]-\infty,+\infty]$ sans aucune restriction sur $(c,q)\in\left(C^0(D^2;\ \mathbb{R}\times\mathbb{R}^d)\right)'$, même pas que $c\geq 0$. On peut en effet montrer que $K(c,q)$ prend toujours la valeur $+\infty$ sauf si : $c\geq 0$, q est absolument continue par rapport à c, de dérivée de Radon-Nikodym v de carré intégrable en c, auquel cas elle prend la valeur

$$\frac{1}{2}\int_{D^2}|v(x,a)|^2 dc(x,a).$$

(On reviendra plus loin sur ces calculs, qu'on pourra admettre pour l'instant, bien qu'il s'agisse d'un exercice élémentaire de théorie de la mesure.) Ainsi on est amené à retranscrire le principe de moindre action entièrement en termes des mesures (c_t,q_t) en leur demandant de minimiser, pour tout intervalle de temps $[t_0,t_1]$ assez petit

$$\int_{t_0}^{t_1} K(c_t,q_t)dt,$$

sous les contraintes que c_t est bistochastique, i.e. $c_t\in DS(D)$, que l'EDP linéaire

$$\partial_t c_t+\nabla_x\cdot q_t=0$$

soit satisfaite et qu'enfin les valeurs aux limites c_{t_0} et c_{t_1} soient fixées dans $DS(D)$. It is, of course, substantially more complicated than the *regular* optimal transport problem, which is related to the Monge-Ampère equation, as discussed in chapter 3.3. However, there are many similarities, in particular the crucial use of convexity tools, as the Fenchel-Rockafellar duality theorem.

4.1 A saddle-point formulation

Using Lagrangian multipliers, this problem can be written as the following "inf-sup" problem (**??**):

$$\begin{cases} \inf_{(X_t)}\sup_{(p_t)} \quad \int_0^T\int_{\mathcal{A}}\{\frac{1}{2}|\frac{d}{dt}X_t(a)|^2-p_t(X_t(a))\}\mu(da)dt \\ \qquad\qquad\qquad +\int_0^T\int_D p_t(x)dx\ dt. \end{cases} \tag{1}$$

This "inf-sup" is obviously bounded from below by the corresponding "sup-inf" problem:

$$\begin{cases} \sup_{(p_t)} \\ \int_0^T\int_D p_t(x)dx\ dt+\inf_{(X_t)}\left[\ \int_0^T\int_{\mathcal{A}}\{\frac{1}{2}|\frac{d}{dt}X_t(a)|^2-p_t(X_t(a))\}\mu(da)dt\right] \end{cases} \tag{2}$$

Notice that this new "relaxed" problem is concave in p. Thus, we can expect it to be more tractable than the original OIT problem and, furthermore, we can hope there is no "relaxation gap", in the sense that, eventually, $\tilde{\inf}\sup=\sup\inf$. Let us look more closely at the "relaxed" problem. Let us fix p and compute the infimum in (X_t). Clearly, a just plays the role of a parameter and we can perform the minimization "a by a". So, the infimum in (X_t) is just

$$\int_{\mathcal{A}} \Psi(T, X_T(a); 0, X_0(a), p)\mu(da),$$

where the "optimal cost" $\Psi(t_1, x_1; t_0, x_0, p)$ between (t_0, x_0) and (t, y) in the field p is defined by

$$\left\{ \begin{array}{c} \Psi(t_1, x_1; t_0, x_0, p) = \\ \\ \inf \left\{ \int_{t_0}^{t_1} \left[\frac{1}{2}|\dot{\eta}(\theta)|^2 - p_\theta(\eta(\theta))\right] d\theta : \ \eta(t_0) = x_0, \eta(t_1) = x_1 \right\}, \end{array} \right. \tag{3}$$

for all $0 \le t_0 \le t_1 \le T, \ \ x_1, x_0 \in D$.
At this level we may use a classical lemma concerning Hamilton-Jacobi equations in the framework of the "weak KAM" theory [REFERENCE fathi]):

Lemma 16. *Let $p(t, x)$ be a sufficiently smooth function on $[0, T] \times D$, where D is a closed, bounded, convex domain in $\mathbb{R}d$. Then*

$$\Psi(t_1, x_1; t_0, x_0, p) = \sup \left\{ \phi(t_1, x_1) - \phi(t_0, x_0) : \quad \partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + p \le 0 \right\}, \tag{4}$$

where Ψ is defined by (3) and the supremum is taken over all C^1 function

$$(t, x) \in [0, T] \times D \to \psi(t, x) \in \mathbb{R}.$$

(For the sake of completeness, a sketch of proof is provided in an appendix on the Hamilton-Jacobi equations at the end of the book.) This suggest to consider the following formulation of the "relaxed" optimal incompressible (ROIT) problem:

$$\left\{ \begin{array}{c} \frac{1}{2T}\tilde{\Delta}^2(X_0, X_T) = \sup_{\varphi, p} J(p, \varphi), \\ \\ J(p, \varphi) = \int_D \int_0^T p_t(x)dt\ dx + \int_d \left[\varphi(T, X_T(a), a) - \varphi(0, X_0(a), a)\right] \mu(da), \\ \\ \text{subject to}: \ \partial_t \varphi(t, x, a) + \frac{1}{2}|\nabla_x \varphi(t, x, a)|^2 + p_t(x) \le 0, \end{array} \right. \tag{5}$$

As a matter of fact, we do not really need lemma 16 and just select (5) as our *new* definition of the relaxed optimal incompressible problem (ROIT). Of course, one of our goals is to establish that there is no difference between $\tilde{\Delta}(X_0, X_T)$ and $\Delta(X_0, X_T)$, as defined in definition (**??**).
$>$From now on, the unknowns are $p_t(x)$ and $\varphi(t, x, a)$ and we will often use the alternative notation $p(t, x)$ for $p_t(x)$. In order to use some results of convex analysis, it is convenient to assume
$\varphi \in C^{1/1/0}_{t/x/a}([0, T] \times D \times \mathcal{A})$, $p \in C^0([0, T] \times \mathcal{A})$ where the notation $C^{1/1/0}_{t/x/a}$ means that we ask C^1-regularity in t and x, and C^0-regularity in a.

Remark 17. Notice that we will not perform any derivation in a, remembering that a should be thought just as a label for each particle in the fluid.

4.2 Rockafellar duality

Now we restate our problem separating the *equality constraints* which will be represented by a functional K_1, and the *inequality constraint* which will be represented by a functional K_2. In other words we want to define convex functionals K_1, K_2 such that

$$\sup_{\varphi,p} J(p,\varphi) = \sup\left[-K_1(A,B) - K_2(A,B)\right]. \tag{6}$$

We will take the functions A, B above to be continuous, and more precisely the second supremum is taken on

$$(A,B) \in C^0(Q,\mathbb{R}) \times C^0(Q,\mathbb{R}d), \text{ where } Q = [0,T] \times D \times \mathcal{A}$$

We then define

$$K_1(A,B) = \begin{cases} -\int_x \int_t p(t,x) - \int_a \left[\varphi(T, X_T(a), a) - \varphi(0, X_0(a), a)\right] \\ \\ \text{if } \exists p \in C^{0/0}_{t/x},\ \varphi \in C^{1/1/0}_{t/x/a} \text{ such that } B = \nabla\varphi,\ A = \partial_t\varphi + p \\ \\ +\infty \text{ else} \end{cases} \tag{7}$$

$$K_2(A,B) = \begin{cases} 0 & \text{if we have the pointwise inequality } A + \frac{1}{2}|B|^2 \leq 0 \text{ on } Q \\ \\ +\infty & \text{otherwise} \end{cases} \tag{8}$$

Remark 18. The definition (7) is consistent, in the sense that if A, B are represented as above by two different couples (φ, p), $(\tilde{\varphi}, \tilde{p})$, then the value of $K_1(A,B)$ is unchanged.

Remark 19. $K_1, K_2 : C^0(Q,\mathbb{R}) \times C^0(Q,\mathbb{R}d) \to \mathbb{R} \cup \{+\infty\}$ are convex functions.

The definitions of K_1, K_2 put us in a setting where it is natural to use the following duality theorem [REFERENCE Brezis][PROOF IN APPENDIX]:

Theorem 1. *[Rockafellar duality theorem] Let E be a real Banach space and consider two functionals $K_1,\ K_2 : E \to \mathbb{R} \cup \{+\infty\}$ which are both convex. Assume that*

$$\textit{there exists a point } u_0 \in E \textit{ such that: } \begin{cases} K_1,\ K_2 \textit{ are finite in } u_0 \\ K_2 \textit{ is continuous in } u_0 \end{cases} \tag{9}$$

then

$$\sup_{u\in E} \left(-K_1(u) - K_2(u)\right) = \inf_{f\in E'} \left(K_1^*(-f) + K_2^*(f)\right), \tag{10}$$

where E' is the dual of E and the Legendre-Fenchel dual $K^* : E' \to \mathbb{R} \cup \{+\infty\}$ of a function $K : E \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$K^*(f) = \sup_{g\in E} \left[\langle f, g\rangle_{E',E} - K(g)\right].$$

Moreover, the infimum in (10) *is reached.*

Remark 20. Surprisingly enough, Rockafellar's duality theorem is quite similar to the Plancherel formula in harmonic analysis. Indeed, at least formally, one can consider the correspondence between the algebraic structures with operations, respectively, $[+, \cdot]$ and $[\max, +]$ (sometimes in this correspondence, inequalities can show up instead of equalities). Then, the Legendre-Fenchel transform is analogous to the Fourier transform and formula (10) corresponds to the Plancherel formula:

$$\int f \cdot g dx = \int \hat{f} \cdot \hat{g} dx.$$

For the reader's convenience, a proof of Rockafellar's theorem is provided in an appendix at the end of the book.

Let us go back to the appolication of Rockafellar's duality theorem, by setting first

$$E = C^0(Q, \mathbb{R}) \times C^0(Q, \mathbb{R}d).$$

Since $D, \mathcal{A}$ (and hence Q) are compact metric spaces, the dual Banach space is the space of Radon measures:

$$E' = \mathcal{M}(Q, \mathbb{R}) \times \mathcal{M}(Q, \mathbb{R}d),$$

and we will denote by $(c, m) \in E'$ a couple of measures in the above decomposition. We will also various notations, depending on the framework, for the duality bracket between $(A, B) \in E$ and $(c, m) \in E'$:

$$\langle (c, m), (A, B) \rangle = \langle c, A \rangle + \langle m, B \rangle = \int_Q (Adc + B \cdot dm)$$

$$= \int_Q A(t, x, a) c(dt, dx, da) + \int_Q B(t, x, a) \cdot m(dt, dx, da)$$

$$= \int_t \int_x \int_a A(t, x, a) c(t, x, a) + \int_t \int_x \int_a B(t, x, a) \cdot m(t, x, a)$$

$$= \int_{t,x,a} Ac + B \cdot m ...$$

Lemma 21. *The functionals $K_1, K_2 : E \to \mathbb{R} \cup \{+\infty\}$ verify the hypotheses of Theorem 1.*

Proof. The convexity condition is clear. Next, we have to find a function u_0 in E having the properties (9). We observe here that there is no chance that K_1 is continuous (for the C^0-norm) because arbitrarily near any function where $K_1 < +\infty$ there is some function with $K_1 = +\infty$. On the other side, in the point $(A_0, B_0) = (-1, 0)$ we have $A_0 = \partial_t \varphi_0 + p_0, B_0 = \nabla \varphi_0$ for $\varphi_0 = 0, p_0 = -1$, so K_1 is finite at this point. On the other side, $K_2(A_0, B_0) = 0$ and this condition is preserved for small perturbations of (A_0, B_0) in the C^0-norm. Therefore the assumptions of Theorem 1 are satisfied. $\square$

We now want to exploit Theorem 1 in our setting. We start by finding a better expression for $K_1^*(-c,-m)$.

Lemma 22.

$$K_1^*(-c,-m) = \sup_{\varphi,p} \Big[-\textstyle\int_{t,x,a} pc + \int_{t,x} p + \int_{t,x,a}(-c\,\partial_t\varphi - m\cdot\nabla\varphi)$$
$$+\int_a (\varphi(T,X_T(a),a) - \varphi(0,X_0(a),a))\Big].$$

This means that $K_1^*(-c,-m)$ takes only values 0 and $+\infty$, with 0 if and only if

$$\begin{cases} \int_{\mathcal{A}} c(t,x,da) = 1, \\ \textit{i.e. } c \textit{ projects to the Lebesgue measure on } [0,T]\times D \textit{ once integrated in } a, \\ \partial_t c + \nabla_x \cdot m = 0, \textit{ with } m(t,\cdot,a) \parallel \partial D \textit{ for all } t,a \\ c(0,x,a) = \delta(x - X_0(a)), \quad c(T,x,a) = \delta(x - X_T(a)). \end{cases}$$

Proof. Using the definitions we find:

$$\begin{aligned} K_1^*(-c,-m) &= \sup\nolimits_{A,B} \Big[\int_{t,x,a}(-Ac - B\cdot m) + \int_{t,x} p \\ &\qquad + \int (\varphi(T,X_T(a),a) - \varphi(0,X_0(a),a))\,da\Big] \\ &= \sup\nolimits_{\varphi,p} \Big[-\int_{t,x,a} pc + \int_{t,x} p + \int_{t,x,a}(-c\,\partial_t\varphi - m\cdot\nabla\varphi) \\ &\qquad + \int_a (\varphi(T,X_T(a),a) - \varphi(0,X_0(a),a))\Big] \end{aligned}$$

Now the functional inside the supremum in the right hand side is linear in (φ,p), so the supremum must be equal to either zero or $+\infty$, and the first case holds only if the functional is equal to zero for all (φ,p). In other words, $K_1^*(-c,-m) < +\infty$ is equivalent to

$$\begin{cases} \text{for all } p \in C^0 \ \int_{t,x,a} pc - \int_{t,x} p = 0 \\ \text{for all } \varphi \in C^{1/1/0}_{t/x/a},\ \int_{t,x,a}(c\partial_t\varphi + m\cdot\nabla\varphi) = \int_a (\varphi(T,X_T(a),a) - \varphi(0,X_0(a),a)) \end{cases}$$

The above conditions are equivalent to the statement of our lemma. $\square$

Next, we pass to the description of

$$K_2^*(c,m) = \sup\left\{\int_{t,a,x} Ac + \int_{t,a,x} B\cdot m\ ;\ A + \frac{1}{2}|B| \le 0\right\}.$$

Lemma 23.

$$K_2^*(c,m) = \begin{cases} \frac{1}{2}\int_{t,a,x} |v|^2 c \text{ if } v = \frac{dm}{dc},\ m \ll c, v \in L^2(Q,dc) \\ \\ +\infty \text{ otherwise.} \end{cases}$$

In this lemma, notation $m \ll c$ means that m is absolutely continuous with respect to c and $\frac{dm}{dc}$ denotes its Radon-Nikodym derivative.

Proof.
- *if c is not positive, then $K_2^*(c,m) = \infty$:* in case this were false, we could indeed find $\tilde{A} \geq 0$ such that $\int \tilde{A}c = -1$ and the integrals defining K_2^* would tend to infinity if applied to the sequence $(A_n, B_n) = (-n\tilde{A}, 0)$.
- *if m is not absolutely continuous with respect to c, then $K_2^*(c,m) = \infty$:* else, there exists $\tilde{B}$ such that $\int |\tilde{B}|c = 0$ and $\int \tilde{B} \cdot m = 1$. It will then suffice to consider the sequence $(A_n, B_n) = \left(-\frac{1}{2}n^2|\tilde{B}|^2, n\tilde{B}\right)$ to prove the claim.

To conclude the proof, it suffices to notice that if $m \ll c$ and $C \geq 0$ then

$$K_2^*(c,m) = \sup\left\{\int_{t,a,x}\left(-\frac{1}{2}|B|^2 + B\cdot v\right)c\right\} = \frac{1}{2}\int_{t,a,x}|v|^2 c, \text{ where } v = \frac{dm}{dc}.$$

□

Finally, we have obtained a rather explicit formulation of the ROIT problem (5):

Theorem 2. *The relaxed optimal incompressible transport (ROIT) problem (5) can be successively written in primal (sup) and dual (inf) form:*

$$\begin{cases} \frac{1}{2T}\tilde{\Delta}^2(X_0, X_T) = \\ \sup_{\varphi,p}\int_D\int_0^T p(t,x)dt\,dx + \int_{\mathcal{A}}\left[\varphi(T, X_T(a), a) - \varphi(0, X_0(a), a)\right]\mu(da), \\ \textit{(subject to : } \partial_t\varphi(t,x,a) + \frac{1}{2}|\nabla_x\varphi(t,x,a)|^2 + p(t,x) \leq 0) \\ = \frac{1}{2}\inf\int_{Q=[0,T]\times D\times\mathcal{A}} |v(t,x,a)|^2 dc(t,x,a) \\ \textit{(subject to : } \partial_t c + \nabla_x\cdot(cv) = 0, \quad cv \parallel \partial D, \quad \int_{\mathcal{A}} c(t,x,da) = 1, \\ c(0,x,a) = \delta(x - X_0(a)), \quad c(T,x,a) = \delta(x - X_T(a))) \end{cases} \tag{11}$$

4.3 Existence and uniqueness of the pressure gradient

We will use the following notation: $f = f(t,x,a)$ (for $a \in \mathcal{A}, x \in D,\ t \in [0,T]$) will always indicate a test function of the regularity implied by one of the measures with which it is in duality. We also introduce a notation for the boundary data:

$$BT(f) = \int_{x,a} \left[f(T,x,a)c(T,x,a) - f(0,x,a)c(0,x,a) \right],$$

where as usual $c(0,x,a) = \delta(x - X_0(a))$, $c(T,x,a) = \delta(x - X_T(a))$.
We consider a minimizer (c, cv) for the dual problem, which exists by Rockafellar's duality theorem, and we denote by *(CE)* the "continuity equation" constraint

$$\forall f \text{ smooth, } BT(f) = \int_{t,x,a} \left[\partial_t f + (v \cdot \nabla) f\right] c,$$

and by *(IC)* the "incompressibililty" constraint $\int_a c = 1$.

Lemma 24. *For all optimal pairs (c, cv), for all pairs $(\tilde{c}, \tilde{v}\tilde{c})$ satisfying (CE) but not necessarily (IC) and for any ε-solution $(p^\varepsilon, \varphi^\varepsilon)$ of the primal problem, we have (with $\int$ meaning $\int_{t,x,a}$)*

$$\int p^\varepsilon (c - \tilde{c}) + \tfrac{1}{2} \int \tilde{c} |\nabla \varphi^\varepsilon - \tilde{v}|^2 + \int \tilde{c} \left| \partial_t \varphi^\varepsilon + \tfrac{1}{2} |\nabla \varphi^\varepsilon|^2 + p^\varepsilon \right|$$

$$\leq \tfrac{1}{2} \int \tilde{c} |\tilde{v}|^2 - \tfrac{1}{2} \int c |v|^2 + \varepsilon^2$$

Proof. We want to use the inequality $\partial_t \varphi^\varepsilon + \frac{1}{2} |\nabla \varphi^\varepsilon|^2 + p^\varepsilon \leq 0$, defining the ε-solutions, together with the fact that $\tilde{c} \geq 0$: we therefore rewrite

$$\begin{aligned} -BT(\varphi^\varepsilon) &= -\int \left(\partial_t \varphi^\varepsilon + (\tilde{v} \cdot \nabla) \varphi^\varepsilon \right) \tilde{c} \\ &= \int \left| \partial_t \varphi^\varepsilon + \frac{1}{2} |\nabla \varphi^\varepsilon|^2 + p^\varepsilon \right| \tilde{c} + \frac{1}{2} \int |\nabla \varphi^\varepsilon - \tilde{v}|^2 \tilde{c} - \frac{1}{2} \int |\tilde{v}|^2 \tilde{c} + \int p^\varepsilon \tilde{c} \end{aligned}$$

By definition of an ε-solution, and since (c, cv) realizes the supremum in the dual problem, we have

$$-BT(\varphi^\varepsilon) - \int p^\varepsilon = -J(p^\varepsilon, \varphi^\varepsilon) \leq -\frac{1}{2} \int |v|^2 c + \varepsilon^2,$$

which inserted in the previous inequality gives the wanted result. □

If in Lemma 24 we take $(\tilde{c}, \tilde{v}) = (c, v)$ we obtain

$$\frac{1}{2} \int c |v - \nabla \varphi^\varepsilon|^2 + \int c \left| \partial_t \varphi^\varepsilon + \frac{1}{2} |\nabla \varphi^\varepsilon|^2 + p^\varepsilon \right| \leq \varepsilon^2. \tag{12}$$

If we were able to pass to the limit in this inequality, we would obtain, as optimality conditions for the ROIT problem:

$$
\begin{cases}
v = \nabla_x \phi, \;\; \partial_t \varphi + \frac{1}{2}|\nabla_x \varphi|^2 + p = 0, \;\; c - \text{a.e.} \;, \\
\partial_t c + \nabla_x \cdot (cv) = 0, \quad cv \parallel \partial D, \;\; \int_{\mathcal{A}} c(t,x,da) = 1, \\
\partial_t \varphi(t,x,a) + \frac{1}{2}|\nabla_x \varphi(t,x,a)|^2 + p(t,x) \leq 0 \;, \forall (t,x,a) \in Q \\
c(0,x,a) = \delta(x - X_0(a)), \quad c(T,x,a) = \delta(x - X_T(a)) \;.
\end{cases}
\tag{13}
$$

Unfortunately, it is unclear that the limit ϕ can be defined in any reasonable sense (this is an open question in the OIT theory). However, we will be shortly able to prove the convergence of ∇p^ε to a definite limit ∇p. To achieve this goal, we first perform smooth deformations of a given pair (c,v) (typically a solution of the dual ROIT problem) into another pair $(\tilde{c},\tilde{v})$ satisfying (CE) but not necessarily (IC). This turns out to be a good way to "feel" how p^ε acts on test functions. We use a definition by duality, requiring that, for all test functions $f(t,x,a) \in \mathbb{R}$ and $B(t,x,a) \in \mathbb{R}d$,

$$
\begin{aligned}
\int_{t,x,a} f\tilde{c} &= \int_{t,x,a} f(t,M(t,x),a)c(t,x,a), \\
\int_{t,x,a} B \cdot (\tilde{c}\tilde{v}) &= \int_{t,x,a} B(t,M(t,x),a) \cdot [(\partial_t + v(t,x,a) \cdot \nabla)M(t,x)]\, c(t,x,a)
\end{aligned}
$$

where $(t,x) \in [0,T] \times D \to M(t,x) \in D$ is a smooth function, so that $M(t,x) = x$ near $\partial([0,T] \times D)$ and $M(t,\cdot)$ is a diffeomorphism of D for all $t \in [0,T]$.

We first observe that under such hypotheses $(\tilde{c},\tilde{v})$ satisfies (CE) as soon as (c,v) satisfies it. Indeed, denoting $\tilde{f}(t,x,a) = f(t,M(t,x),a)$, we find:

$$
\int \left[\partial_t \tilde{f} + (\tilde{v} \cdot \nabla)\tilde{f}\right] \tilde{c} = \int ((\partial_t f)(t,M(t,x),a)
$$

$$
+(\nabla_x f)(t,M(t,x,a) \cdot [\partial_t M + v \cdot \nabla]M(t,x))c(t,x,a)
$$

$$
= \int \left[\partial_t \tilde{f} + v \cdot \nabla_x \tilde{f}\right] c = BT(f),
$$

where we have used (CE) for (c,v) and the chain rule for $\tilde{f}$.

Now, let us rewrite the conclusion of Lemma 24 where $(\tilde{c},\tilde{v})$ is as above. We first treat the term:

$$
\int p^\varepsilon \tilde{c} = \int p^\varepsilon(t,M(t,x))c(t,x,a) = \int p^\varepsilon(t,M(t,x))dtdx,
$$

where we used the (IC) condition for c.
Next, we write

$$\frac{1}{2}\int \tilde{c}|\tilde{v}|^2 = \sup_{A+\frac{1}{2}|B|^2\le 0}\int A(t,x,a)\tilde{c}+B(t,x,a)\cdot\tilde{m} = \sup_B \int \left(-\frac{1}{2}|B|^2 + B\cdot\tilde{v}\right)\tilde{c}$$

$$= \sup_B \int [-\frac{1}{2}|B(t,M(t,x),a)|^2 + B(t,M(t,x),a)\cdot(\partial_t + v\cdot\nabla)M(t,x)]c$$

$$= \sup_{\tilde{B}} \int \left[-\frac{1}{2}|\tilde{B}|^2 + \tilde{B}\cdot(\partial_t + v\cdot\nabla)M)\right] c$$

$$= \frac{1}{2}\int |(\partial_t M(t,x) + v(t,x,a)\cdot\nabla)M(t,x)|^2\, c(t,x,a),$$

where $\tilde{B}(t,x,a) = B(t,M(t,x),a)$.

The inequality in Lemma 24 now reads

$$\int_{t,x}(\tilde{p}^\varepsilon - p^\varepsilon) + \int_{t,x,a} c|\widetilde{\partial_t\varphi^\varepsilon} + \tfrac{1}{2}|\widetilde{\nabla\varphi^\varepsilon}|^2 + \tilde{p}^\varepsilon| + \tfrac{1}{2}\int_{t,x,a}|\widetilde{\nabla\varphi^\varepsilon} - \partial_t M - (v\cdot\nabla)M|^2\, c$$
$$\le \tfrac{1}{2}\int_{t,x,a} c|\partial_t M + (v\cdot\nabla)M|^2 - \tfrac{1}{2}\int_{t,x,a} c|v|^2 + \varepsilon^2, \tag{14}$$

where we always denote $\tilde{f}(t,x,a) = f(t,M(t,x),a)$.

Remark 25. Inequality (14) is more tractable than (24), since here the dependence on $(\tilde{c},\tilde{v})$ is substituted by the dependence on the smooth function M.

Let us now use the following variant of "Moser's Lemma" [MOSER, DACOROGNA-MOSER, RIVIERE-YE]

Lemma 26 (Moser's Lemma for $\mathbb{T}^d$)**.** *Let $\sigma_0,\sigma_1 \in C^\infty(\mathbb{T}^d)$ be strictly positive probability densities (i.e. $\sigma_i > 0$, $\int_{\mathbb{T}^d}\sigma_i dx = 1$ for $i = 0,1$). Then there exists a diffeomorphism $M : \mathbb{T}^d \to \mathbb{T}^d$ with $OPERAdet(DM) > 0$ such that for all continuous test functions φ there holds*

$$\int_{\mathbb{T}^d}\varphi(M(x))\sigma_0(x)dx = \int_{\mathbb{T}^d}\varphi(x)\sigma_1(x)dx.$$

Proof. We will find an expression of M as the flow $N(t,x)$ at time $t = 1$ of a vectorfield $z(t,x)$:

$$\begin{cases} \partial_t N(t,x) = z(t,N(t,x)) \\ N(0,x) = x \end{cases}$$

To impose the right conditions on z, we express the pushforward density obtained from $\sigma_0(x)dx$ via $N(t,\cdot)$:

$$\int \varphi(N(t,x))\sigma_0(x)dx = \int \varphi(x)\sigma(t,x)dx \text{ for all } t,\ \varphi \in C^\infty(\mathbb{T}^d)$$

The flow equation then gives us the evolution equation $\partial_t \sigma + \nabla \cdot (z\sigma) = 0$ for $\sigma(t,x)$. If we ask that $\sigma(t,x) = (1-t)\sigma_0(x) + t\sigma_1(x)$, then the above equation assumes a much simpler form: $(\sigma_1 - \sigma_0)(x) = -\nabla \cdot [\sigma(t,x)z(t,x)] = -\nabla \cdot Z(x)$. We make the extra Ansatz that $Z = \nabla\zeta$, and we obtain the equation

$$\Delta\zeta + \sigma_1 - \sigma_2 = 0 \text{ on } \mathbb{T}^d.$$

The integrability condition for this equation is $\int(\sigma_1 - \sigma_0) = 0$, which is satisfied in our case. Therefore we obtain a smooth solution ζ. The vectorfield z can now be expressed in terms of $\zeta, \sigma_0, \sigma_1$ and it is bounded because of the strict positivity condition on σ_0, σ_1:

$$z(t,x) = \frac{\nabla\zeta(x)}{(1-t)\sigma_0(x) + t\sigma_1(x)},$$

and since z is smooth and bounded, also N is smooth, therefore $M(x) = N(1,x)$ is smooth, as wanted. □

Remark 27.
- For this version of Moser's Lemma, we needed σ_0, σ_1 to be strictly positive.
- In [Dacorogna moser] this kind of lemma is done on a compact domain $D \subset \mathbb{R}d$ and is followed by a second step where the boundary condition $M(x) = x$ on ∂D is ensured. This somehow hints at the fact that the possible constructions are more flexible, and that the results could be ameliorated. This is done in [Riviere Ye].

We will need the following refinement of Moser's Lemma:

Lemma 28. *Let $\theta \in C_c^\infty(]0,T[)$ be a nonnegative function and $w \in C_c^\infty(OPERAintD, \mathbb{R}d)$. If $||\theta||_{L^\infty}$ is small enough, we can find a family of diffeomorphisms $M(t,x)$ such that $M(t,x) = x$ near $\partial([0,T] \times D)$ and for all $\varphi \in C_c^1(\mathbb{R}d)$there holds*

$$\int_D \varphi(M(t,x))dx = \int_D \varphi(x)dx + \theta(t)\int_D \nabla\varphi(x) \cdot w(x)dx.$$

Moreover M will be representable as a flow, i.e. there will hold

$$\partial_t M(t,x) = z(t, M(t,x)),$$

where $z(t,x) = \frac{\theta'(t)w(x)}{1-\theta(t)[\nabla \cdot w(x)]}$.

Proof. Call $S = ||\theta||_{L^\infty}$, so that $\theta([0,T]) = [0,S]$. We observe that since θ has compact support, $\theta(0) = 0$. We start by defining

$$\begin{aligned} \tilde{\sigma}(s,x) &= 1 - s\nabla \cdot w(x) \\ \tilde{w}(s,t) &= \frac{w(x)}{\tilde{\sigma}(s,x)}, \end{aligned}$$

so that $\partial_s \tilde\sigma + \nabla \cdot (\tilde w \tilde\sigma) = 0$. We then consider the flow of $\tilde w$. We define

$$\begin{cases} \partial_s \tilde M(s,x) = \tilde w(s, \tilde M(s,x)) \text{ for } s \in [0,S] \\ \tilde M(0,x) = x \end{cases}$$

Then clearly $\tilde M(s,x) = x$ for x near ∂D. We observe that $\tilde\sigma(0,x) = 1$ and that for all $\varphi \in C^0(D)$

$$\int \varphi(\tilde M(s,x))dx = \int \varphi(x)\tilde\sigma(s,x)dx.$$

We then define $M(t,x) = \tilde M(\theta(t) - \theta(0), x) = \tilde M(\theta(t), x)$, and we have

$$\begin{aligned} \partial_t M(t,x) &= \partial_t \tilde M(\theta(t),x) = \tilde w(\theta(t), M(t,x))\theta'(t) \\ &= \frac{w(M(t,x))}{\tilde\sigma(\theta(t), M(t,x))}\theta'(t) \\ &= \frac{w(M(t,x))}{1 - \theta(t)\nabla \cdot w(M(t,x))}\theta'(t) \\ &= z(t, M(t,x)) \end{aligned}$$

We can also compute

$$\begin{aligned} \int \varphi(M(t,x))dx &= \int \varphi(\tilde M(\theta(t),x))dx \\ &= \int \varphi(x)\tilde\sigma(\theta,x)dx \\ &= \int \varphi(x)dx - \theta(t)\int \varphi(x)(\nabla \cdot w(x))dx \\ &= \int \varphi(x)dx + \theta(t)\int \nabla\varphi(x)\cdot w(x)dx, \end{aligned}$$

as wanted. $\square$

Now we can use the above Lemma to rewrite the pressure integrals in (14):

$$\int [p^\varepsilon(t, M(t,x)) - p(t,x)]dx = \theta(t)\int \nabla p^\varepsilon(t,x)\cdot w(x)dx, \tag{15}$$

and by (14), we get for ∇p^ε, viewed as a distribution on $]0,T[\times \mathring{D}$,

$$\langle \nabla p^\varepsilon, \theta \otimes \omega\rangle = \int_{t,x} \nabla p^\varepsilon(t,x)\theta(t)\cdot w(x) \le \varepsilon^2 + \frac{1}{2}\int_{t,x,a} \left(|\partial_t M + v\cdot\nabla M|^2 - |v|^2\right)c. \tag{16}$$

Thus, we see that, as a distribution, ∇p^ε is bounded on $]0,T[\times \mathring{D}$ uniformly in ε. Up to a subsequence we then have $\nabla p^\varepsilon \rightharpoonup \nabla p$ in the sense of distributions, combining Banach-Steinhaus and Banach-Alaoglu theorems.

Uniqueness of the limit ∇p

We consider again the inequality (16) concerning the pressure, but we take a limit in the time-dependent test function $\theta(t)$ more carefully:

$$\theta(t) = \delta\zeta(t)\text{for } \zeta \in C_c^\infty(]0,T[), \text{ and for } |\delta| \text{ small}$$

therefore $M(t,x) = \delta\zeta(t)w(x)$. We now want to take the limit in (16) as $\delta \to 0$. Therefore we start by computing:

$$\begin{aligned} M(t,x) - x &= O(\delta) \\ \partial_t M(t,x) &= \delta\zeta'(t)w(x) + O(\delta^2) \\ M(t,x) &= x + \delta\zeta(t)w(x) + O(\delta^2) \\ \frac{\partial}{\partial x_j}M(t,x) &= \delta_{ij} + \delta\zeta(t)\frac{\partial w}{\partial x_j}(x) + O(\delta^2), \end{aligned}$$

and inserting this in the integrand in the right hand side of (16), we obtain

$$\begin{aligned} |\partial_t M + v \cdot \nabla M|^2 - |v|^2 &= \frac{1}{2}\left(\left|\delta\zeta'(t)w_j(x) + v_i + \sum_j v_j \delta\zeta(t)\partial_j w_i + O(\delta^2)\right|^2 - |v|^2\right) \\ &= \sum_i \delta\left[\zeta'(t)w_i(x) + \sum_j v_j\zeta\partial_j w_i\right] v_i + O(\delta^2), \end{aligned}$$

and since (16) should hold along the subsequence $\varepsilon_n \to 0$ such that $\nabla p^{\varepsilon_n} \rightharpoonup \nabla p$ found in the previous section and for all δ small enough, we obtain (first passing $n \to \infty$ then $\delta \to 0$)

$$\begin{aligned} \langle \nabla p, \theta \otimes w\rangle &= \sum_i \int_{t,x,a}\left[\zeta' w_i + \sum_j v_j \partial_j w_i \zeta\right] c v_i \\ &= -\sum_i \langle \partial_t \int_a c v_i + \sum_j \partial_j \int_a c v_i v_j, \zeta \otimes w_i\rangle, \end{aligned}$$

which means that in the sense of distributions,

$$-\nabla p = -\partial_t \int_a cv + \nabla \cdot \int_a cv \otimes v. \tag{17}$$

Since this is true for *every* optimal solution (c, cv), ∇p is uniquely defined. This means that the limit ∇p is unique as a distribution, and in particular it does not depend on the sequence ∇p^{ε_n} which we choose. Therefore $\nabla p^\varepsilon \to \nabla p$.

Remark 29 (regularity of the pressure field). From the above discussion we obtain that ∇p is the derivative of a measure. By working substantially harder, in [REF BRENIER 99] ∇p was shown to be itself a measure, and an improvement on the time regularity was achieved in [REF AmbrFigalli], where $\nabla p \in L^2(]0,T[, BV(\mathring{D}))$ was shown. The best one can expect would be that $D_x^2 p$ is a measure, which is still a conjecture. Better regularity however cannot be achieved, due to the existence of irregular examples [REF NEXT LESSONS].

4.4 Convergence of approximate solutions

Definition 30. *We say that a couple* $(c^\varepsilon, m^\varepsilon) \in E'$ *(where we recall that* $E = C^0(Q; \mathbb{R} \times \mathbb{R}d)$*,* $Q = [0, T] \times D \times \mathcal{A}$*) is an* approximate solution *if:*
i) $m^\varepsilon \ll c^\varepsilon$*,* $m^\varepsilon = c^\varepsilon v^\varepsilon$*;*
ii) the continuity equation and the incompressibility constraint (in the sense of distributions) hold in the limit $\varepsilon \to 0$ *(we denote them by, respectively,* (ACE) *and* (AIC)*);*
ii) $2K(c^\varepsilon, m^\varepsilon) = \int_{t,a,x} |v^\varepsilon|^2 c^\varepsilon \to T^{-1} \tilde{\Delta}^2(X_0, X_T)$ *as* $\varepsilon \to 0$*.*

Theorem 3. *There is a unique pressure gradient* ∇p *which enjoys the property that all approximate solutions* $(c^\varepsilon, m^\varepsilon)$*, in the sense of definition (30) satisfy*

$$\partial_t \int_a c^\varepsilon v^\varepsilon + \nabla_x \cdot \int_a c^\varepsilon v^\varepsilon \otimes v^\varepsilon \to -\nabla p, \tag{18}$$

as $\varepsilon \to 0$*, in the sense of distributions. This pressure gradient is precisely the one just found in the study of the ROIT problem.*

Proof. We first observe that we have compactness of the positive measures c^ε:

$$\int_{t,x,a} c^\varepsilon \to \int_{[0,T]\times D} dxdt =< \infty.$$

For the measures $|m^\varepsilon|$ we get

$$\int_{t,x,a} |m^\varepsilon| \le \sqrt{\int \frac{|m^\varepsilon|^2}{c_\varepsilon}} \sqrt{\int c^\varepsilon} = \sqrt{K(c^\varepsilon, m^\varepsilon)} \sqrt{\int c^\varepsilon} \to \frac{1}{\sqrt{2}T} \tilde{\Delta}(X_0, X_T).$$

>From the above two boundedness results it follows that up to extracting a subsequence we may assume that $(c^\varepsilon, m^\varepsilon)$ converge to a measure (c, m) weakly. Passing to the limit in the equations (ACE) and (AIC) we obtain (CE), (IC). Moreover, by lower semicontinuity (looking at K in the dual formulation [REF 2lesson] we obtain $K(c, m) \le \liminf K(c^\varepsilon, m^\varepsilon) = \frac{1}{2T} \tilde{\Delta}^2(X_0, X_T)$, but since (c, m) is an admissible generalized flow, we obtain that the equality should hold, therefore (c, m) is optimal.
Now, we show the convergence of $\int_a c^\varepsilon v^\varepsilon \otimes v^\varepsilon$ to $\int_a cv \otimes v$ in the sense of distributions. To do this we first observe that by compactness, there exist a symmetric-matrix valued measure $\nu(t, x, a)$ and a subsequence $\varepsilon_n \to 0$ such that

$$\int_a c^{\varepsilon_n} v^{\varepsilon_n} \otimes v^{\varepsilon_n} \to \nu \text{ weakly.}$$

Then by lower semicontinuity we have $\nu \ge \int_a cv \otimes v = \tilde{\nu}$ in the sense of symmetric-matrix valued measures. But since we already know that

$$\int_{t,x} OPERAtr(\nu) = \int_{t,a,x} c^{\varepsilon_n} |v^{\varepsilon_n}|^2 \to K(c, m) = \int_{t,a} OPERAtr(\tilde{\nu}),$$

We obtain $\nu = \tilde{\nu}$. Since for optimal pairs like (c, m) we know that, in the sense of distributions,

$$\partial_t \int_a cv + \nabla \cdot \int_a cv \otimes v = \nabla p,$$

and since ∇p is unique, by the above compactness results we obtain that any subsequence $\varepsilon_k \to 0$ has a subsequence satisfying (18), for a uniquely determined pressure field p. This completes the proof. $\square$

4.5 Shnirelman's approximate solutions

A crucial ingredient to close the gap between the optimum of the relaxed optimal incompressible problem $\frac{1}{2}\tilde{\Delta}^2(X_0, X_1)$ and $\frac{1}{2}\Delta^2(X_0, X_1)$, as introduced in definition (**??**), is the following rephrasing of a result due to Shnirelman [REFERENCE shnirelman]

Theorem 4 (Shnirelman's approximation theorem)**.** *Let D be a bounded convex domain in $\mathbb{R}d$ with $d \geq 2$. Let $(c, m) \in E'$ be an admissible solution to the ROIT problem, i.e. satisfying (IC) and (CE) conditions with $K(c, m) < +\infty$. Then, we can find, for every small $\varepsilon > 0$, a smooth divergence-free vector field $v(t, x)$, compactly supported in the interior of $[0, T] \times D$, with associated volume-preserving flow $g_t(x)$, defined by*

$$\frac{d}{dt} g_t(x) = v(t, g_t(x)), \quad g_0(x) = x,$$

such that

$$\begin{cases} \int_{\mathcal{A}} |X_T(a) - g_T(X_0(a))|^2 \mu(da) \leq \varepsilon^2, \\ \\ \frac{1}{2} \int_{t,x} |v(t, x)|^2 \leq K(c, m) + \varepsilon^2. \end{cases}$$

In addition, from this result, we immediately obtain approximate solutions as in Definition 30, by setting:

$$\begin{cases} X^\varepsilon(t, a) = g_t^\varepsilon(X_0(a) \\ c^\varepsilon(t, x, a) = \delta(x - X^\varepsilon(t, a)) \\ m^\varepsilon(t, x, a) = \partial_t X^\varepsilon(t, a) c^\varepsilon(t, x, a) = v^\varepsilon(t, X^\varepsilon(t, a)) c^\varepsilon(t, x, a) = v^\varepsilon(t, x) c^\varepsilon(t, x, a) \end{cases}$$

We then can verify (ACE):

$$\begin{aligned} \int_{t,x,a} [\partial_t f + v^\varepsilon \cdot \nabla f]\, c^\varepsilon = \int_{t,a} [\partial_t f(t, X^\varepsilon(t, a), a) + \partial_t X^\varepsilon(t, a) \cdot (\nabla_x f)(t, X^\varepsilon(t, a), a)] \\ = \int_a [f(T, X^\varepsilon(T, a), a) - f(0, X^\varepsilon(0, a), a)] \\ \to \int_a f(T, X_T(a), a) - \int_a f(0, X_0(a), a) = \langle c_T, f(T, \cdot, \cdot) \rangle - \langle c_0, f(0, \cdot, \cdot) \rangle, \end{aligned}$$

as wanted. As for the verification of (AIC), we have:

$$\int_{t,x,a} f(t,x)c^\varepsilon(t,x,a) = \int_{t,a} f(t,X^\varepsilon(t,a))$$

$$= \int_{t,a} f(t,g_t^\varepsilon(X_0(a)) = \int_{t,x} f(t,g_t^\varepsilon(x)) = \int_{t,x} f(t,x),$$

since X_0 has the Lebesgue measure for law on D and g_t^ε is volume preserving. Finally, we verify the convergence of the energy:

$$\begin{aligned}
K(c^\varepsilon, m^\varepsilon) &= \inf_{A+\frac{1}{2}|B|^2\leq 0} \int_{t,a,x} [Ac + B\cdot m] \\
&= \inf_{A+\frac{1}{2}|B|^2\leq 0} \int_{t,a} [A(t,X^\varepsilon(t,a),a) + \partial_t X^\varepsilon(t,a)\cdot B(t,X^\varepsilon(t,a),a)] \\
&= \frac{1}{2}\int_{a,t} |\partial_t X^\varepsilon(t,a)|^2 \\
&= \frac{1}{2}\int_{a,t} |v^\varepsilon(t,X^\varepsilon(t,a))|^2 = \frac{1}{2}\int_{x,t} |v^\varepsilon(t,x)|^2 \\
&\to \frac{1}{2T}\tilde{\Delta}^2(X_0,X_T),
\end{aligned}$$

Finally, the existence these "Shnirelman" approximate solutions, combined with the convergence theorem 3, first show that there is no gap between the ROIT optimum $\tilde{\Delta}^2(X_0,X_T)$ and $\Delta^2(X_0,X_T)$, as defined in definition (**??**), and finally leads to theorem **??**.

For the sake of completeness we provide as the end of the book, in an appendix, a rather explicit version of Shnirelman's theorem, when $D=[0,1]^2$, $d\geq 2$, for admissible solutions (c,m) such that $m\cdot e=0$ for one of the direction e of the cube, say $e=(0,\cdot,\cdot,\cdot,0,1)$. Let us call them "flat" admissible solutions. This flatness property allows us to play with the last coordinate to construct, rather explicitely, the desired vector field v which, in general, needs a tiny but non-trivial component $e\cdot v$ to do the approximation correctly. As a matter of fact, the flatness condition is sufficient to cover all data X_0, X_T that "coincide" in direction e, namely: $e\cdot(X_T(a)-X_0(a))=0$. This is precisely for this kind of data that Shnirelman was able to prove the non-existence of classical solutions to the OIT problem [SHNIRELMAN]. Therefore, the flatness condition is perfectly relevent for the mathematical theory. In addition, from the physical point of view, the flatness condition is directly related to the popular "hydrostatic approximation" of the Euler equations used in geo-sciences to describe fluid motions in thin domains, such as lakes, oceans or the atmosphere.

The mathematical core of this chapter on optimal incompressible transport.is now closed. However, we would like, in the next subsections, to give a more concrete approach to the OIT problem, by providing several examples of more or less explicit solutions, including some solutions related to the "hydrostatic approximation" we just mentioned.

4.6 Consistency of the Euler equations with the optimal incompressible transport problem

Let us first prove, as a consistency result, that sufficiently smooth solutions to the Euler equations, under some conditions, are solutions to the OIT problem. Going back to definition (**??**) of the Euler equations, we consider a time interval $[0, T]$ sufficiently short so that

$$T^2 \left(D_x^2 p(t,x)\right)[\xi,\xi] \le \alpha \pi^2 |\xi|^2, \quad \forall (t,x,\xi) \in [0,T] \times D \times \mathbb{R}d. \tag{19}$$

where $\alpha \in [0, 1]$ is a fixed constant. We start with the elementary observation.

Lemma 31. *Assume that D is a bounded closed convex domain in $\mathbb{R}d$. Under assumption (19), given $x, y \in D$, any solution $\eta(t) \in D$ of the second-order differential equation*

$$\frac{d^2}{dt^2}\eta(t) + (\nabla p_t)(\eta(t)) = 0, \tag{20}$$

is a minimizer of

$$I_p(\eta) = \int_0^T \left[\frac{1}{2}|\frac{d}{dt}\eta(t)|^2 - p(t,\eta(t))\right] dt$$

in the class $\mathcal{C} = \{\eta : [0,T] \to D : \ \eta(0) = y, \ \eta(T) = x\}$. In addition, if $\alpha < 1$, then η is the unique minimizer.

Proof. We compare the values $I_p(\eta)$ and $I_p(\tilde\eta)$ for a "competitor" $\tilde\eta \in \mathcal{C}$. Our claim is that:

$$\int_0^T \left[\frac{1}{2}|\frac{d}{dt}\eta(t)|^2 - p(t,\eta(t))\right] dt \le \int_0^T \left[\frac{1}{2}|\frac{d}{dt}\tilde\eta(t)|^2 - p(t,\tilde\eta(t))\right] dt, \tag{21}$$

with strict inequality if $\alpha < 1$ and $\tilde\eta \neq \eta$. By Taylor expansion, we have

$$p(t,\tilde\eta(t)) \le p(t,\eta(t)) + \nabla p(t,\eta(t)) \cdot [\tilde\eta(t) - \eta(t)] + \frac{1}{2}K(p)|\tilde\eta(t) - \eta(t)|^2, \tag{22}$$

where, since D is convex and we are assuming (19), we can take

$$K(p) = \sup_{x,t} |D_x^2 p(t,x)| \le \frac{\pi^2}{T^2}. \tag{23}$$

The number π enters assumption (19) because of the Poincaré inequality:

$$\int_0^T |\tilde\eta(t) - \eta(t)|^2 dt \le \frac{T^2}{\pi^2}\int_0^T |\frac{d}{dt}\tilde\eta(t) - \frac{d}{dt}\eta(t)|^2 dt, \tag{24}$$

that can be proven by trigonometric expansions:

$$(\eta - \tilde\eta)(t) = \sum_{n\in\mathbb{N}} a_n \sin(n\pi t/T), \quad (a_n \in \mathbb{R}d, \ \ n \in \mathbb{N})$$

>From the above three inequalities (22), (23) and (24) we get (forgetting in our notations the dependencies on the integrating variable t),

$$\int_0^T [p(\tilde{\eta}) - p(\eta) - \nabla_x p(\eta) \cdot (\tilde{\eta} - \eta)]\, dt \leq \frac{1}{2} \int_0^T |\frac{d}{dt}\tilde{\eta} - \frac{d}{dt}\eta|^2 dt.$$

We now use the equation (20), which in the above condensed notations reads $\ddot{\eta} = -\nabla_x p(\eta)$, and we integrate by parts. We obtain

$$\int_0^T \left[p(\tilde{\eta}) - p(\eta) - \frac{d}{dt}\eta \cdot (\frac{d}{dt}\tilde{\eta} - \frac{d}{dt}\eta)\right] dt \leq \frac{1}{2} \int_0^T |\frac{d}{dt}\tilde{\eta} - \frac{d}{dt}\eta|^2 dt,$$

which can be rewritten as (21). Notice that the uniqueness also follows from $\alpha < 1$. $\square$

Remark 32. The convexity of D was only needed on order to have the estimate (23) directly from the assumption (19) and to be able to include it in a Taylor formula. The same strategy of proof as in Lemma 31 works for any connected domain D such that its diameter as a length metric subspace of $\mathbb{R}d$ is finite, once modified our assumption (19) by a factor proportional to such diameter.

Let us now move back to a solution $(X_t(a), p(t,x))$ of the Euler equation (**??**). We want to compare p to any other potential $\tilde{p}(t,x)$ given on $[0,T] \times D$. We consider the "dual" OIT functional

$$J(\tilde{p}) = \int_0^T dt \int_{\mathcal{A}} \tilde{p}(t,x)dx + \int_{\mathcal{A}} I_{\tilde{p}}(X_0(a), X_T(a))\mu(da), \tag{25}$$

where

$$I_{\tilde{p}}(y,z) = \inf \left\{ \int_0^T \left[\frac{1}{2}|\frac{d}{dt}\eta(t)|^2 - \tilde{p}(t,\eta(t))\right] dt :\ \eta(t) \in D,\ \eta(0) = y,\ \eta(T) = z \right\}.$$

Theorem 5. *Let $(X_t(a), p(t,x))$ be a solution of the Euler equation (**??**) satisfying condition (19) on the time interval $[0,T]$. Then p is a maximizer of the dual IOT functional (25).*

Proof. By definition we have for any a

$$I_{\tilde{p}}(X_0(a), X_T(a)) \leq \int_0^T \left[\frac{1}{2}|\frac{d}{dt}X_t(a)|^2 - \tilde{p}(t, X_t(a))\right] dt,$$

so we can write

$$\begin{aligned} J(\tilde{p}) &\leq \int_0^T \int_D \tilde{p}(t,x)dxdt + \int_0^T \int_{\mathcal{A}} \left[\frac{1}{2}|\frac{d}{dt}X_t(a)|^2 - \tilde{p}(t, X_t(a))\right] \mu(da) \\ &= \int_0^T \int_D p(t,x)dxdt + \int_0^T \int_{\mathcal{A}} \left[\frac{1}{2}|\frac{d}{dt}X_t(a)|^2 - p(t, X_t(a))\right] \mu(da) \\ &= \int_0^T \int_D p(t,x)dxdt + \int_{\mathcal{A}} I_p(X_0(a), X_T(a))\mu(da) \\ &= J(p). \end{aligned}$$

Here we have used that X_t has the Lebesgue measure as law on D for the first equality, and Lemma 31 for the second one. More precisely we have used

$$\int_D \tilde{p}(t,x) - \int_{\mathcal{A}} \tilde{p}(t,X_t(a))\mu(da) = 0 = \int_D p(t,x) - \int_{\mathcal{A}} p(t,X_t(a))\mu(da)$$

and the fact that, for each a, $t \to X_t(a)$ is minimizing $I_p(X_0(a), X_T(a))$. $\square$

4.7 Hydrostatic solutions to the Euler equations

Definition (**??**) that we have used for the Euler equations, written in terms of trajectories, is much more flexible than the classical one (**??**), written in term of vector fields. Let us consider for example a "classical" solution $(v(t,x), p(t,x))$ of the Euler equations, in the sense of definition (**??**), in a very flat three-dimensional domain such as $D_\varepsilon = D \times [0,\varepsilon]$, where D is a bounded convex closed domain in $\mathbb{R}2$. Let us rescale the vertical coordinate x^3 and the third component v^3 of the velocity field: $(x^3, v^3) \to (\varepsilon x^3, \varepsilon v^3)$. After this rescaling we get, on the rescaled 3D domain $D \times [0,1]$, no longer the Euler equations but a rescaled version of them, namely

$$I_\varepsilon D_t v + \nabla p = 0, \quad D_t = \partial_t + v \cdot \nabla, \quad \nabla \cdot v = 0, \tag{26}$$

where I_ε denotes the diagonal matrix $I_\varepsilon = \mathrm{diag}(1,1,\varepsilon^2)$. Notice that the operators D_t and $\nabla\cdot$ are unchanged and ε only features in I_ε. It is very customary in geo-sciences to neglect ε by substituting $I_0 = \mathrm{diag}(1,1,0)$ for I_ε. This is the so-called hydrostatic approximation, for which the pressure does not depend on the vertical coordinate x^3:

$$I_0 D_t v + \nabla p = 0, \quad D_t = \partial_t + v \cdot \nabla, \quad \nabla \cdot v = 0. \tag{27}$$

This approximation of the 3D Euler equations in thin domain is very common in ocean-atmosphere computational models. As an evolution equation, the hydrostatic limit of the Euler equations is much more singular than the original Euler equations: it is ill-posed, in some sense, on any linear Sobolev space, but well-posed on some adequate functional convex cone [BRENIER, GRENIER, MASMOUDI]). Of course, all smooth solutions of the 2D Euler equations on the 2D domain D are particular solutions of this hydrostatic limit, but there are many other solutions that are genuinely three dimensional.
Let us consider a smooth solution $(v(t,x), p(t,x))$ of the hydrostatic limit (27) of the Euler equations.on the 3D domain $D \times [0,1]$. We denote by $g_t(x)$ the volume-preserving flow in $D \times [0,1]$ generated by

$$\frac{d}{dt} g_t(x) = v(t, g_t(x)), \quad g_0(x) = x, \quad x \in D \times [0,1]$$

and define $\overline{X}_t(a) = g_t(\overline{X}_0(a)) \in D$, for all $a \in \mathcal{A}$, where $\overline{X}_0$ is a given Borel map $A \to D \times [0,1]$ with the 3D Lebesgue measure on $D \times [0,1]$ as law. Let us now denote $X_t(a) = (\overline{X}_t^1(a), \overline{X}_t^2(a)) \in D$ the two first components of $\overline{X}_t(a)$.

This defines a time-dependent family of maps $A \to D$ having the 2D Lebesgue measure on D as law. Indeed, if we consider a continuous function f on $\mathbb{R}2$, we can trivially lift it as a continuous function F on $\mathbb{R}3$ by setting $F(x^1, x^2, x^3) = f(x^1, x^2)$ and we get

$$\int_{\mathcal{A}} f(X_t(a))\mu(da) = \int_{\mathcal{A}} F(\overline{X}_t(a))\mu(da) = \int_{D\times[0,1]} F(x^1, x^2, x^3)dx^1dx^2dx^3$$

$$= \int_{D\times[0,1]} f(x^1, x^2)dx^1dx^2dx^3 = \int_D f(x^1, x^2)dx^1dx^2,$$

which is enough to show that X_t has the 2D Lebesgue measure on D as law. Meanwhile, since (v, p) solves the hydrostatic limit of the Euler equations (27), we get for X:

$$\frac{d^2}{dt^2}X_t(a) + (\nabla p)(t, X_t(a)) = 0$$

where ∇ denotes the two-dimensional gradient on the two-dimensional domain D. (We again have used that $p(t, x)$ does not depend on x^3 and, therefore, can be seen as a time-dependent function on the two-dimensional domain D.) So, we have obtained that $(X_t(a), p(t, x))$ is a solution of the Euler equations on the 2D domain D, in the sense of definition (**??**), although it does not correspond, in general, to a solution of the 2D Euler equations in the classical sense of (**??**).
Even more provocative is the perspective of 1D solutions to the Euler equations. Indeed, in the classical setting of formulation (**??**), there are only trivial solutions of the Euler equations. On a bounded domain suchas $[0, 1]$, the only possible solution is $v = 0$. However, there are many non-trivial 1D solutions to the Euler equations with definition (**??**). Once again, such solutions can be obtained by rescaling a thin 2D domain and by passing to the hydrostatic limit in the 2D Euler equations, by dimension reduction, exactly as we did from three to two dimensions.
At this stage, one may wonder whether there is a reciprocal to this construction. More precisely, givem a solution to the Euler equation on a $d-$dimensional domain D, in the general sense of (**??**), can we see this solution as an hydrostatic limit of a "conventional" solution to the $d+1$ dimensional Euler equations, in the sense of (**??**), on a thin $d+1$ domain $D \times [0, \varepsilon]$, as ε goes to zero? The answer is formally yes, but there are analytic difficulties in doing it properly. Analytically, the problem is as follows. We are given a smooth solution $(X_t(a), p(t, x))$ to the Euler equation (**??**) on some domain D in $\mathbb{R}d$ and we want to recover X_t as the d first component of a map $\overline{X_t} \in D \times [0, 1]$ carried by a $d + 1$ divergence-free vector field $\overline{v}(t, \overline{x})$ defined for $\overline{x} \in D \times [0, 1]$ and solution to the hydrostatic limit of the $d + 1$ Euler equations, namely

$$I_0 D_t\overline{v} + \overline{\nabla}p = 0, \quad D_t = \partial_t + \overline{v} \cdot \overline{\nabla}, \quad \overline{\nabla} \cdot \overline{v} = 0, \tag{28}$$

where $\overline{\nabla}$ is the $d + 1$ dimensional gradient on $D \times [0, 1]$, $I_0 = \mathrm{diag}(1, \cdot, \cdot, \cdot, 1, 0)$, and $p(t, \overline{x}) = p(t, x)$ is unchanged, depending only on t and the d first space coordinates x of $\overline{x}$.

4.8 Explicit solutions to the OIT problem

The simplest non trivial explicit 1D solution to the Euler equations in the sense of (**??**) known to us can be written as follows. We take $D = [-1, 1]$ and define the label space $(\mathcal{A}\mu)$

$$\mathcal{A} = \{a = (x, \omega) \in D \times [0, 1]\}$$

equipped with the Lebesgue measure. We set

$$X_t(x, \omega) = x \cos t + \sqrt{1 - x^2}\ \sin t \cos(2\pi\omega), \quad p(t, x) = p(x) = x^2/2 \tag{29}$$

and can check (easily) that $\frac{d^2}{dt^2} X_t(x, \omega) = -X_t(x, \omega)$ and (not so easily) that the law of X_t on D is the Lebesgue measure. Observe that, for $T = \pi$, we $X_0(x, \omega) = x$ and $X_T(x, \omega) = -x$, while $p''(x) = 1$. Thus, for these data, $(X_t(a), p(t, x))$ is solution to the IOT problem according to Theorem 5, since condition (19) is *sharply* satisfied.

A related example can be defined in 2D on the unit disk D (with normalized Lebesgue measure). The formulae are very similar. (Actually the previous 1D solution is more or less a projection of this one.) The label space is

$$\mathcal{A} = \{a = (x, \omega) = (x^1, x^2, \omega) \in D \times [0, 1]\}$$

equipped with the (normalized) Lebesgue measure and we set

$$X_t(x, \omega) = x \cos t + \sqrt{1 - |x|^2}\ \sin t \exp(2\pi i\omega), \quad p(t, x) = |x|^2/2 \tag{30}$$

with an abusive complex notation. Observe that, in both case, we have $D_x^2 p(t, x) = Id$, $X_0(x, \omega) = x$, $X_T(x, \omega) = -x$, if we choose $T = \pi$. This proves that both solutions are solutions to the OIT problem, according to Theorem 5, since, once again, condition (19) is sharply satisfied. This OIT amounts to transfering all particles from their initial position to the opposite one on the unit disk D, during the time interval $[0, \pi]$, in an incmpressible fashion inside D. Of course the obtained motion is not at all conventional: every "particle" issued from x in the unit disk get split according to the "microscopical" (or "hiddem") variable ω and follow a continuum of different trajectories parameterized by $\omega \in [0, 1]$, with equal probability, and eventually reaches its destination $-x$ at time $T = \pi$. This strange motion looks much more conventional, once lifted as a 3D incompressible motion by adding a vertical coordinate x^3 along a small interval of length ε, and projecting back to the 2D basis. This is just another example of hydrostatic limit of the 3D Euler equation. The multiplicity of trajectories observed on the 2D domain D just correspond to the projection of three dimensional trajectories in $D \times [0, \varepsilon]$. Accordingly, the "hidden" variable ω is just keeping record (in a non-trivial way) of the missing vertical coordinate x^3.

It is interesting to notice, that in the 2D case, there are two other solutions X^+ and X^- to the same OIT problem, namely

$$X_t^+(x, \omega) = x \exp(it), \quad X_t^-(x, \omega) = x \exp(-it), \quad p(t, x) = |x|^2/2, \tag{31}$$

with an obvious complex notation. They actually do not depend on the "micro" variable ω and correspond to two classical solutions of the 2D Euler equations with (stationary) velocity fields $v^+(x) = (-x^2, x^1)$, $v^-(x) = (x^2, -x^1)$. Geometrically, they correspond to simple rigid rotations of the unique disk. We further point out that these three different solutions to the same IOT problem share the same pressure field, which is fully consistent with Theorem **??**. Surprinsingly enough, there is a very rich family of other solutions to the same OIT problem, obtained by M. Bernot, A.Figalli and F. Santambrogio [BERNOT-FIGALLI-SANTAMBROGIO]. In particular, solution (30) can be "decomposed" into two more "fundamental" solutions of the Euler equations in the sense of (**??**) (which was very surprizing to us).

5 Hamilton-Jacobi equations

5.1 A tribute to Fourier and Gauss!

So far, we have so much emphasised the interest of convex methods that we have entirely omitted the role of Fourier methods in PDEs ! This section can be seen as an *intermezzo* and a tribute to Fourier, paradoxically devoted to the Hamilton-Jacobi equation

$$\text{(HJ)} \qquad \partial_t \phi + \frac{1}{2}|\nabla\phi|^2 = 0,$$

which is a rare example of PDE for which, not only the Fourier analysis, but also the Lebesgue spaces can be entirely ignored, in particular thanks to the remarkable theory of so-called "viscosity solutions", by Crandall, Evans and Lions [102], which relies only on the concept of continuous and semi-continuous functions, without any reference to Lebesgue spaces and, of course, to the Fourier analysis. A typical result is the full justification of the "Hopf formula" which provides the unique solution $\phi(t,x)$ of the HJ equation in terms of its initial data $\phi(0,x)$, for all $t \geq 0$ and $x \in \mathbb{R}^d$, through:

$$\phi(t,x) = \inf_{\xi\in\mathbb{R}^d} \ \frac{|\xi - x|^2}{2t} + \phi(0,\xi).$$

This formula is very much related to convex analysis (and more specifically to the Legendre-Fenchel transform). The purpose of this section is to show how everything can be deduced from the heat equation and the way Fourier solved it.

The basic idea comes from Feynman's interpretation of Quantum Mechanics with his concept of "path integrals". However, let us start at a more conventional level by reminding the well known solution of the heat equation thanks to Gaussian integrals that follow almost instantaneously from its Fourier analysis. More precisely,

on introduit le "semi-groupe de la chaleur" sur l'espace euclidien $\mathbb{R}^d$

$$(S_\epsilon(t)v)(x) = \int_{\mathbb{R}^d} \exp(-\pi|y|^2)v(x + \sqrt{2\pi\epsilon t}\, y)dy, \quad t \geq 0, \quad x \in \mathbb{R}^d.$$

Rappelons que

$$\int_{\mathbb{R}^d} \exp(-\pi|y|^2)dy = 1$$

et $S_\epsilon(t)$ s'interprête donc, pour chaque $t \geq 0$ et $\epsilon > 0$, (comme un opérateur de "moyenne locale" autour de chaque point x. Il est facile de voir que cet opérateur linéaire est bien défini sur tous les espaces $L^p(\mathbb{R}^d)$, pour $p \in [1, +\infty]$ et, avec un peu de calcul, qu'il est contractant au sens large

$$||S_\epsilon(t)v||_{L^p} \leq ||v||_{L^p} \ , \quad \forall t, s \geq 0.$$

(On pourra utiliser l'inégalité de Jensen pour le vérifier. Noter les cas $p = 1$ et $p = +\infty$ qui sont particulièrement simples.) On parle de semi-groupe car on peut vérifier, en utilisant le fait que les fonctions gaussiennes sont stables par convolution, que :

$$S_\epsilon(0) = I_d, \quad S_\epsilon(t+s) = S_\epsilon(t) \circ S_\epsilon(s), \quad \forall t, s \geq 0.$$

(Il ne s'agit en revanche pas d'un groupe et $S(t)$ n'est pas défini pour $t < 0$). Enfin, on parle de "chaleur", car on peut calculer que $u(t, x) = (S_\epsilon(t)v)(x)$, qui est de classe C^∞ en $(t, x) \in]0, +\infty[\times\mathbb{R}^d$, est solution, au sens classique, de l'équation de la chaleur $\partial_t u = \epsilon\Delta u/2$, avec, pour "donnée initiale" $u(0, \cdot) = v$, puisque $S(0) = I_d$.

[Notons qu'on obtient une formule analogue pour l'équation de Schrödinger

$$\partial_t u = -\frac{\epsilon}{2i}\Delta u$$

(qui a des propriétés très différentes de celle de la chaleur) en posant

$$(\mathfrak{S}_\epsilon(t)v)(x) = \int_{\mathbb{R}^d} \exp(i\pi|y|^2)v(x + \sqrt{2\pi\epsilon t}\, y)dy, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^d,$$

ce qui représente une intégrale oscillante à valeurs complexes et définit alors $t \in \mathbb{R} \to \mathfrak{S}_\epsilon(t)$ comme un groupe d'isométries de $L^2(\mathbb{R}^d)$.]

5.2 Lien (rapide !) avec le mouvement brownien

Fixons $T > 0$. Par la propriété de semi-groupe, on a pour tout $N \geq 1$

$$S_\epsilon(T) = S_\epsilon(T/N)^N$$

ce qui permet d'écrire

$$(S_\epsilon(T)v)(x) = \int_{(\mathbb{R}^d)^N} \exp(-\pi \sum_{k=1}^{N} |y_k|^2)v\left(x + \sum_{k=1}^{N} y_k\sqrt{\frac{2\pi\epsilon T}{N}}\right) dy_1 \cdots dy_N,$$

ce qu'on peut interpréter comme l'espérance de la valeur de v au point atteint après une marche aléatoire, partie de x, faite de N pas indépendants y_k dans $\mathbb{R}^d$, $k = 1, \cdots, N$, de loi gaussienne centrée, ϵ mesurant le niveau du bruit.

[On peut d'ailleurs étudier le passage à la limite $N \to \infty$ ce qui est une façon de définir le mouvement brownien et la mesure de Wiener, puis finalement d'écrire la solution $u(T, x)$ de l'équation de la chaleur comme espérance de v évaluée au bout d'une trajectoire brownienne issue de x au temps 0 et parvenue au temps T.]

5.2.1 Transformation de Hopf, lemme de Laplace et équation d'Hamilton-Jacobi

Le semi-groupe S_ϵ conserve manifestement la positivité, et, de façon un tout petit peu moins évidente, vérifie $u_\epsilon(t, x) = (S_\epsilon(t)v)(x) > 0$, dès que $t > 0$ et à condition que $v \geq 0$ ne soit pas presque partout nulle. Il est alors pertinent d'écrire u_ϵ sous la forme exponentielle suivante

$$u_\epsilon(t, x) = \exp(-\frac{\phi_\epsilon(t, x)}{\epsilon}).$$

De

$$\partial_t u_\epsilon = \epsilon \Delta u_\epsilon / 2,$$

on tire assez facilement que

$$\partial_t \phi_\epsilon + \frac{1}{2} |\nabla \phi_\epsilon|^2 = \epsilon \Delta \phi_\epsilon / 2.$$

[En effet

$$\epsilon du_\epsilon / u_\epsilon = \epsilon d(\log u_\epsilon) = -d\phi_\epsilon$$

d'où

$$\epsilon \partial_t u_\epsilon = -u_\epsilon \partial_t \phi_\epsilon \ ,$$

$$\epsilon \nabla u_\epsilon = -u_\epsilon \nabla \phi_\epsilon \ ,$$

$$\epsilon \Delta u_\epsilon = \epsilon^{-1} u_\epsilon |\nabla \phi_\epsilon|^2 - u_\epsilon \Delta \phi_\epsilon \ ,$$

et finalement

$$0 = -\epsilon \partial_t u_\epsilon + \epsilon^2 \Delta u_\epsilon / 2 = u_\epsilon |\nabla \phi_\epsilon|^2 / 2 - \epsilon u_\epsilon \Delta \phi_\epsilon / 2 + u_\epsilon \partial_t \phi_\epsilon,$$

comme prévu, en divisant par u_ϵ.]
Cela nous donne espoir de résoudre l'équation complètement non-linéaire, dite d'Hamilton-Jacobi

$$\partial_t \phi + \frac{1}{2} |\nabla \phi|^2 = 0,$$

en passant à la limite $\epsilon \to 0$!
C'est ce qu'on appelle parfois la limite de "viscosité évanescente" qui nous permet de résoudre une équation "du premier ordre" comme limite d'une équation parabolique "du deuxième ordre" (en espace). Cette idée a été poursuivie par E. Hopf en 1950, dans le cas d'une variable d'espace. Notons que, pour $\epsilon > 0$, le gradient de ϕ_ϵ, $B = \nabla\phi_\epsilon$ est solution du système du deuxième ordre "quasi-linéaire":

$$\partial_t B + \nabla(\frac{|B|^2}{2}) = \epsilon\Delta B/2,$$

qui s'appelle habituellement équation de Burgers en dimension 1 ($d = 1$)

$$\partial_t B + \partial_x(B^2/2) = \epsilon(\partial_x)^2 B/2, \quad x \in \mathbb{R}.$$

(Il semblerait qu'il faudrait plutôt l'appeler équation de...Bateman !) Essayons maintenant de passer à la limite dans la formule donnant explicitement

$$u_\epsilon(t,x) = \exp(-\frac{\phi_\epsilon(t,x)}{\epsilon})$$

comme solution de l'équation de la chaleur, à savoir, comme on l'a vu plus haut,

$$u_\epsilon(t,x) = \int_{\mathbb{R}^d} \exp(-\pi|y|^2) v(x + \sqrt{2\pi\epsilon t}\, y) dy, \quad t \geq 0, \quad x \in \mathbb{R}^d.$$

On fait dépendre la donnée initiale de ϵ et on l'écrit sous la forme :

$$v(x) = v_\epsilon(x) = \exp(-\frac{\psi(x)}{\epsilon})$$

de sorte que $\psi(x)$, qui elle ne dépend pas de ϵ, s'interprête comme la donnée initiale de $\phi_\epsilon(t,x)$ et on suppose ψ uniformément continue. On suppose aussi ψ indépendante de ϵ et de plus que

$$\lim_{|x|\to\infty} \frac{|\psi(x)|}{1+|x|^2} = 0.$$

On a donc

$$u_\epsilon(t,x) = (2\pi\epsilon t)^{-d/2} \int_{\mathbb{R}^d} \exp(-\frac{|\xi - x|^2}{2\epsilon t}) v_\epsilon(\xi) d\xi$$

(en effectuant le changement de variable $y \to \xi = x + \sqrt{2\pi\epsilon t}\, y$)

$$= (2\pi\epsilon t)^{-d/2} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{\epsilon}\left(\frac{|\xi - x|^2}{2t} + \psi(\xi)\right)\right) d\xi.$$

Comme

$$u_\epsilon(t,x) = \exp(-\frac{\phi_\epsilon(t,x)}{\epsilon}),$$

on a

$$\phi_\epsilon(t,x) = -\epsilon \log u_\epsilon(t,x) = \log(2\pi\epsilon t)\epsilon d/2 - \epsilon \log \int_{\mathbb{R}^d} \exp\left(\frac{1}{\epsilon} F(\xi; t, x)\right) d\xi,$$

où

$$F(\xi; t, x) = -\frac{|\xi - x|^2}{2t} - \psi(\xi).$$

C'est alors qu'on utilise le fameux "lemme de Laplace" (qui est au départ de la théorie des "grandes déviations" en théorie des probabilités):

Lemma 33. *Soit A un ensemble Lebesgue mesurable non négligeable dans $\mathbb{R}^d$ et une fonction F Lebesgue mesurable, telle que*

$$0 < \int_A \exp(F(\xi)) d\xi < +\infty$$

alors, lorsque $\epsilon \downarrow 0$,

$$\epsilon \log \left(\int_A \exp(\frac{F(\xi)}{\epsilon}) d\xi \right) \to \sup \mathrm{ess}_A \ F.$$

Preuve

On prend d'abord ϵ sous la forme

$$\epsilon = \frac{1}{1+R}$$

de sorte que

$$I = \int_A \exp(\frac{F(\xi)}{\epsilon}) d\xi = \int_A \exp(F(\xi)) \exp(RF(\xi)) d\xi.$$

On note L le sup essentiel de F sur A et

$$J = \int_A \exp(F(\xi)) d\xi.$$

On a $J > 0$ et $J < +\infty$ par hypothèse. Le log de J est donc fini.
On commence par la majoration évidente

$$I \leq \exp(RL) \int_A \exp(F(\xi)) d\xi$$

et donc

$$\epsilon \log I = \frac{1}{R+1} \log I \leq \frac{1}{R+1} (RL + \log J) \to L, \quad \epsilon \downarrow 0.$$

Pour la minoration de I, on considère $\lambda < L$, quelconque. Par définition de L comme sup essentiel de F sur A, on peut trouver un sous ensemble mesurable B de A de mesure de Lebesgue positive tel que $F(\xi) \geq \lambda$ pour tout $\xi \in B$. On a forcément

$$K = \int_B \exp(F(\xi)) d\xi \in]0, +\infty[.$$

[En effet K est plus petit que I et donc fini. Par ailleurs, $K \geq \exp(\lambda) \int_B d\xi > 0$.]
On a donc

$$I \geq \int_B \exp(F(\xi)) \exp(RF(\xi)) d\xi \geq \exp(R\lambda) \int_B \exp(F(\xi)) d\xi = \exp(R\lambda) K$$

et donc

$$\epsilon \log I = \frac{1}{R+1} \log I \geq \frac{1}{R+1} (R\lambda + \log K) \to \lambda, \quad \epsilon \downarrow 0.$$

Comme λ peut être choisi arbitrairement proche de L, le lemme est démontré.

Fin de la preuve.

Appliquons maintenant le lemme de Laplace, pour tout $t > 0$ et tout x fixés, à la solution $\phi_\epsilon(t, x)$ de l'équation de Hamilton-Jacobi "parabolisée"

$$\partial_t \phi_\epsilon + \frac{1}{2} |\nabla \phi_\epsilon|^2 = \epsilon \Delta \phi_\epsilon / 2$$

avec donnée initiale $\phi_\epsilon(0, \cdot) = \psi$, qui ne dépend pas de ϵ. On rappelle que

$$\phi_\epsilon(t, x) = \log(2\pi\epsilon t)\epsilon d/2 - \epsilon \log \int_{\mathbb{R}^d} \exp\left(\frac{1}{\epsilon} F(\xi; t, x)\right) d\xi,$$

où

$$F(\xi; t, x) = -\frac{|\xi - x|^2}{2t} - \psi(\xi).$$

Comme on a supposé

$$\lim_{|\xi| \to \infty} \frac{|\psi(\xi)|}{1 + |\xi|^2} = 0,$$

les hypothèses du lemme sont satisfaites en posant $A = \mathbb{R}^d$ et $F(\xi) = F(\xi; t, x)$ (avec abus de notation, (t, x) étant fixé). On obtient donc, à la limite,

$$\phi(t, x) = \inf_{\xi \in \mathbb{R}^d} \frac{|\xi - x|^2}{2t} + \psi(\xi)$$

ce qui fournit une formule naturelle, dite de Hopf, pour la solution de l'équation d'Hamilton-Jacobi

$$\partial_t \phi + \frac{1}{2} |\nabla \phi|^2 = 0.$$

[On aurait pu deviner cette formule par le raisonnement fallacieux, quoiqu'instructif, suivant: on écrit l'équation sous la forme

$$\sup_{w \in \mathbb{R}^d} \left(\partial_t \phi + w \cdot \nabla \phi - \frac{|w|^2}{2}\right) = 0$$

(jusqu'ici rien d'incorrect) et on fait le pari (a priori injustifié) que

$$\phi(t, x) = \inf_{w \in \mathbb{R}^d} \Phi(t, x; w)$$

où la fonction "génératrice" Φ est solution de l'équation linéaire sous-jacente, à coefficient constant en (t, x), $w \in \mathbb{R}^d$ étant vu comme un paramètre,

$$\partial_t \Phi + w \cdot \nabla \Phi - \frac{|w|^2}{2} = 0,$$

avec comme donnée initiale $\Phi(0, x; w) = \psi(x)$. On trouve (quasi) immédiatement

$$\Phi(t, x; w) = \Phi(0, x - tw, w) + t\frac{|w|^2}{2} = \psi(x - tw) + t\frac{|w|^2}{2},$$

ce qui nous donnerait

$$\phi(t, x) = \inf_{w \in \mathbb{R}^d} \psi(x - tw) + t\frac{|w|^2}{2} = \inf_{\xi \in \mathbb{R}^d} \frac{|\xi - x|^2}{2t} + \psi(\xi),$$

ce qui est bien la formule de Hopf. Bien qu'injustifié, ce raisonnement permet pourtant de deviner ou de retrouver rapidement le résultat !]

On parle naturellement de "solution de viscosité" (sous-entendu : "évanescente"). Notons que la solution fournie par la formule de Hopf est de régularité limitée quand t grandit, à moins que ψ soit convexe et lisse. En effet, dans le cas contraire, dès que t devient assez grand, pour x fixé, l'inf peut être atteint en plusieurs points ξ distincts, ce qui ruine la régularité de ϕ en (t, x), quelle que soit la régularité de ψ. On a là une propriété très caractéristique des équations d'évolution du premier ordre, la perte de régularité en temps fini de leurs solutions, ce qui rend leur étude particulièrement ardue.

5.2.2 La théorie de Crandall-Evans-Lions des solutions de viscosité

Il est tentant d'aller au delà de la formule assez miraculeuse de Hopf pour traiter des équations non-linéaires plus générales, typiquement,

$$\partial_t \phi + H(t, x, \nabla \phi) = 0.$$

voire

$$\partial_t \phi + H(t, x, \nabla \phi, D^2 \phi) = 0,$$

en supposant que la fonction $(t, x, w, M) \to H(t, x, w, M) \in \mathbb{R}$
i) soit suffisamment lisse par rapport à toutes les variables (respectivement $t \in \mathbb{R}_+$, $x \in \mathbb{R}^d$, $w \in \mathbb{R}^d$ et M dans l'ensemble des matrices symétriques $d \times d$);
ii) admette des bornes du type

$$|H(t, x, w, M)| \leq C(1 + |w|^\alpha + |M|^\beta)$$

pour des constantes C, α, β convenables,
iii) soit monotone décroissante par rapport à la dernière variable

(au sens que $H(t,x,w,M) \leq H(t,x,w,\tilde{M})$ dès que $\tilde{M} - M$ est une matrice symétrique positive au sens large).

C'est tout l'objet de la théorie initiée par Crandall, Evans et Lions (P.-L.) [102] au début des années 1980, sous le nom de théorie des "solutions de viscosité", d'abord dans le cas des équations du premier ordre de type Hamilton-Jacobi

$$\partial_t \phi + H(t,x,\nabla\phi) = 0.$$

Une solution de viscosité ϕ, supposée a priori seulement continue (voire Lipschitz mais en tous cas pas C^1), y est définie d'une façon particulièrement originale et ingénieuse. On se donne une fonction test $\zeta(t,x)$ et on considère n'importe quel point (t_0,x_0) où $\phi - \zeta$ atteint son minimum (ou, plus généralement, un minimum local). Si ϕ était aussi lisse que ζ, on déduirait donc que

$$\partial_t \phi(t_0,x_0) = \partial_t \zeta(t_0,x_0), \quad \nabla\phi(t_0,x_0) = \nabla\zeta(t_0,x_0),$$

ce qui nous suggère de remplacer, au point (t_0,x_0), les dérivées de ϕ (qui sont typiquement mal ou pas définies) par celles de ζ. Ainsi on demandera que

$$\partial_t \zeta(t_0,x_0) + H(t_0,x_0,\nabla\zeta(t_0,x_0)) \geq 0.$$

Pour les points de maximum, on demandera l'inégalité dans l'autre sens :

$$\partial_t \zeta(t_0,x_0) + H(t_0,x_0,\nabla\zeta(t_0,x_0)) \leq 0.$$

[Avec un peu d'imagination, on verra là une sorte de version $(\max,+)$ des formulations faibles au sens des distributions !]
Il est très remarquable que cette théorie, contrairement à celle des distributions, ne fait pas appel à la théorie de l'intégration de Lebesgue et aux espaces L^p, mais seulement à celle des fonctions continues. Elle aurait pu donc être découverte bien avant Lebesgue, mais il a fallu attendre les années 1980 pour qu'elle apparaisse (ce qui fournit un thème de réflexion sur la chronologie d'éventuelles mathématiques extra-terrestres:-).

5.2.3 La formule de Feynman-Kac, le principe de moindre action et la programmation dynamique

On sait (probablement depuis Euler!) que, pour des matrices carrées $m \times m$ A et B on a

$$\exp(t(A+B)) = \lim_{N\uparrow\infty} (\exp(tA/N)\exp(tB/N))^N, \quad \forall t \geq 0.$$

On va étendre (formellement) cette formule (dite de Lie-Trotter) aux deux opérateur linéaires en dimension infinie:
i) le semi groupe de la chaleur (qu'on a introduit au début du chapitre)

$$S_\epsilon(t) = \exp(\epsilon t\Delta/2)$$

(avec une notation symbolique très parlante)
ii) l'opérateur de multiplication (plus trivial) $T_\epsilon(t) = \exp(t\Phi/\epsilon)$ défini par :

$$(T_\epsilon(t)v)(x) = v(x)\exp(t\Phi(x)/\epsilon),$$

où la fonction $\Phi : \ x \in \mathbb{R}^d \to \Phi(x) \in \mathbb{R}$ (appelée "potentiel") est donnée (avec des hypothèses convenables de régularité et de comportement à l'infini qu'on ne précisera pas ici). En appliquant, formellement, la formule de Lie-Trotter, on trouve

$$\exp\left(t(\epsilon\Delta/2 + \Phi/\epsilon)\right) = \lim_{N\uparrow\infty} \left(S_\epsilon(t/N)T_\epsilon(t/N)\right)^N,$$

ce qui s'interprète en disant que l'équation parabolique

$$\partial_t u = (\epsilon\Delta/2 + \Phi/\epsilon)u = \epsilon\Delta u/2 + u\Phi/\epsilon$$

admet comme solution $u = u(t,x)$, de donnée initiale u_0,

$$u(t,x) = \lim_{N\uparrow\infty} \left((S_\epsilon(t/N)T_\epsilon(t/N))^N u_0\right)(x).$$

(En fait, cette formule se justifie complètement avec un peu d'analyse fonctionnelle).
Plus explicitement, on voit que

$$(S_\epsilon(t/N)T_\epsilon(t/N))v)(x) = (2\pi\epsilon t/N)^{-d/2}\int_{\mathbb{R}^d} \exp(\frac{t\Phi(x_0)}{\epsilon N} - N\frac{|x-x_0|^2}{2\epsilon t})v(x_0)dx_0$$

$$= (2\pi\epsilon t/N)^{-d/2}\int_{\mathbb{R}^d} \exp(\epsilon^{-1}t/N\left(\Phi(x_0) - \frac{1}{2}|\frac{x-x_0}{t/N}|^2\right))v(x_0)dx_0$$

et donc, itérativement,

$$\forall t \geq 0, \quad \forall x_N \in \mathbb{R}^d, \quad (S_\epsilon(t/N)T_\epsilon(t/N))^N v)(x_N)$$

$$= (2\pi\epsilon t/N)^{-Nd/2}\int_{(\mathbb{R}^d)^N} \exp(\epsilon^{-1}\frac{t}{N}\sum_{k=0}^{N-1}\left(\Phi(x_k) - \frac{1}{2}|\frac{x_{k+1}-x_k}{t/N}|^2\right))v(x_0)dx_0 \cdots dx_{N-1}.$$

On a envie de passer à la limite $N \uparrow \infty$ et de conjecturer que l'équation

$$\partial_t u = \epsilon \Delta u/2 + u\Phi/\epsilon,$$

avec donnée initiale u_0, admet pour solution :

$$u(t,x) = \int_{\Omega_{t,x}} \exp(\epsilon^{-1} \int_0^t \left(\Phi(\xi_s) - \frac{1}{2}|\frac{d\xi_s}{ds}|^2\right) ds)\ u_0(\xi_0)\text{"}d\xi\text{"}$$

où $\Omega_{x,t} \subset (\mathbb{R}^d)^{[0,t]}$ serait l'ensemble des chemins $s \in [0,t] \to \xi_s \in \mathbb{R}^d$ tels que $\xi_t = x$ et "$d\xi$" serait une sorte de mesure de Lebesgue sur cet ensemble ! C'est la fameuse formule de Feynman-Kac, qui, telle quelle, n'a pas de sens (vu l'impossibilité avérée de construire la mesure de Lebesgue sur le produit infini $(\mathbb{R}^d)^{[0,t]}$), et il faut en fait introduire la mesure de Wiener, associée au mouvement brownien, pour la rendre rigoureuse (ce qui dépasse le cadre de ce cours).

[En fait l'idée de cette formule remonte à la thèse de Feymann où l'équation de Schrödinger

$$i\epsilon\partial_t u = (-\epsilon^2\Delta/2 + \Phi)u$$

est "résolue" par la célèbre formule "de Feynman"

$$u(t,x) = \int_{\Omega_{t,x}} \exp(-i\epsilon^{-1} \int_0^t \left(\Phi(\xi_s) - \frac{1}{2}|\frac{d\xi_s}{ds}|^2\right) ds)\ u_0(\xi_0)\text{"}d\xi\text{"}$$

qui ne peut en aucun cas être justifiée dans le cadre de la théorie de la mesure usuelle, contrairement à celle de Feynman-Kac, pour laquelle la mesure de Wiener suffit. Elle est néanmoins le point de départ d'une part essentielle de la physique théorique moderne : QED, modèle standard, théorie des cordes, etc...]

Une application, tout à fait formelle et injustifiée à ce stade, du lemme de Laplace à la formule de Feynman-Kac permet même de conjecturer que, dans la limite $\epsilon \downarrow 0$, on trouve pour l'équation d'Hamilton-Jacobi

$$\partial_t\phi + \frac{|\nabla\phi|^2}{2} + \Phi = 0,$$

la solution suivante, avec donnée initiale ϕ_0,

$$\phi(t,x) = \inf_{\xi\in\Omega_{t,x}} \phi_0(\xi_0) + \int_0^t \left(\frac{1}{2}|\frac{d\xi_s}{ds}|^2 - \Phi(\xi_s)\right) ds,$$

dite "formule de programmation dynamique" qui peut complètement se justifier dans le cadre des solutions de viscosité à la Crandall-Evans-Lions, mais, là encore, cela dépasse les limites de ce cours. En fait la formule reste valable quand le potentiel Φ dépend aussi de la variable t et on rapprochera ce résultat du principe de moindre action évoqué dans un précédent chapitre.

6 Panov formulation of scalar conservation laws

6.1 A short review of first order systems of conservation laws and Kruzhkov's theory of scalar conservation laws

First order systems of conservation laws read:

$$\partial_t u + \sum_{i=1}^{d} \partial_{x_i}(Q_i(u)) = 0,$$

or, in short, using the nabla notation,

$$\partial_t u + \nabla \cdot (Q(u)) = 0,$$

where $u = u(t,x) \in \mathbb{R}^m$ depends on $t \geq 0$, $x \in \mathbb{R}^d$, and $\cdot$ denotes the inner product in $\mathbb{R}^d$. The Q_i (for $i = 1, \cdots, d$) are given smooth functions from $\mathbb{R}^m$ into itself. The system is called hyperbolic when, for each $\tau \in \mathbb{R}^d$ and each $U \in \mathbb{R}^m$, the $m \times m$ matrix $\sum_{i=1,d} \tau_i Q_i'(U)$ can be put in diagonal form with real eigenvalues. There is no general theory to solve globally in time the initial value problem for such systems of PDEs. (See [32, 106, 187, 213] for a general introduction to the field.) In general, smooth solutions are known to exist for short times but are expected to become discontinuous in finite time. Therefore, it is usual to consider discontinuous weak solutions, satisfying additional "entropy conditions", to adress the initial value problem. Some special situations are far better understood. First, for some very special (but nevertheless very important in Physics and Geometry) systems (enjoying "linear degeneracy" or "null conditions"), smooth solutions may be global (shock free), at least for "small" initial data (see [161, 174, 222], for instance). This includes the famous result on the stability of the Minkowski space in General Realivity by Klainerman and Christodoulou [100]. Next, in one space dimension $d = 1$, for a large class of systems, existence and uniqueness of global weak entropy solutions have been proven Bianchini et Bressan for initial data of sufficiently small total variation [31]. Still, in one space dimension, for a limited class of systems (typically for $m = 2$), existence of global weak entropy solutions have been obtained for large initial data by "compensated compactness" arguments [225, 114, 177]. Finally, there is a very comprehensive theory in the much simpler case of a single "scalar" conservation laws, i.e. when $m = 1$. Kruzhkov [166] showed that such a scalar conservation law has a unique "entropy solution" $u \in L^\infty$ for each given initial condition $u_0 \in L^\infty$. (If the derivative Q' is further assumed to be bounded, then we can substitute L^1_{loc} for L^∞ in this statement.) An entropy (or Kruzhkov) solution is an L^∞ function that satisfies the following distributional inequality

$$\partial_t C(u) + \nabla_x \cdot (Q^C(u)) \leq 0,$$

for all Lipschitz convex function $C : \mathbb{R} \to \mathbb{R}$, where the derivative of Q^C is defined by $(Q^C)' = C'Q'$ (the initial condition u_0 being prescribed by continuity at $t = 0$, in L^1_{loc}, namely:

$$\lim_{t \to 0} \int_B |u(t,x) - u_0(x)| dx = 0,$$

for all compact subset B of $\mathbb{R}^d$). Beyond their existence and uniqueness, the Kruzhkov solutions enjoy many interesting properties. Each entropy solution $u(t, \cdot)$, with initial condition u_0, continuously depends on $t \geq 0$ in L^1_{loc} and can be written $T(t)u_0$, where $(T(t), t \geq 0)$ is a family of order preserving operators:

$$T(t)u_0 \ \geq \ T(t)\tilde{u}_0 \ , \ \ \forall t \geq 0,$$

whenever $u_0 \geq \tilde{u}_0$. Since constants are trivial entropy solutions to a scalar conservation law, it follows that if u_0 takes its values in some fixed compact interval, so does $u(t, \cdot)$ for all $t \geq 0$. Next, two solutions u and $\tilde{u}$, with $u_0 - \tilde{u}_0 \in L^1$, are L^1 stable with respect to their initial conditions:

$$\int |u(t,x) - \tilde{u}(t,x)| dx \leq \int |u_0(x) - \tilde{u}_0(x)| dx,$$

for all $t \geq 0$. As a consequence, the total variation $TV(u(t, \cdot))$ of a Kruzhkov solution u at time $t \geq 0$ cannot be larger than the total variation of its initial condition u_0. This easily comes from the translation invariance of the scalar conservation law and from one of the most classical definitions of the total variation of a function v, namely:

$$TV(v) = \sup_{\eta \in \mathbb{R}^d, \ \eta \neq 0} \int \frac{|v(x+\eta) - v(x)|}{|\eta|} dx,$$

where $|\cdot|$ denotes the Euclidean norm on both $\mathbb{R}$ and $\mathbb{R}^d$. As a matter of fact, the space L^1 plays a key role in Kruzhkov's theory. Indeed, there is no L^p stability with respect to initial conditions in any $p > 1$. Typically, for $p > 1$, the Sobolev norm $||u(t, \cdot)||_{W^{1,p}}$ of a Kruzhkov solution blows up in finite time. This fact has induced a great amount of pessimism about the possibility of a unified theory of global solutions for general multidimensional systems of hyperbolic conservation laws. Indeed, simple linear systems, such as the wave equation (written as a first order system) or the Maxwell equations, are not well posed in any L^p but for $p = 2$ [80]. However, as we are going to see that L^2 turns out to be a perfectly suitable space for entropy solutions to multidimensional scalar conservation laws, provided a different formulation is used, based on a combination of level-set, kinetic and transport-collapse approximations, in the spirit of previous works by Giga, Miyakawa, Osher, Tsai and the author [36, 38, 39, 52, 148, 228]. As a matter of fact, this new formulation was already obtained by Panov [203] and just rediscovered, in a different style, by the author in [58]. Let us also mention the more recent approach of Serre and Vasseur where the space L^2 can also be used for conservation laws, from a quite different angle [216]. Finally let us emphasise that this new formulation à la Panov is entirely convex, and provides a remarkable example of "hidden convexity" in nonlinear PDEs.

6.2 The main result

N.B. For notational simplicity, we limit ourself to initial conditions u_0 that can be written as

$$u_0(x) = \int_0^1 1\{Y_0(a,x) < 1/2\}da,$$

for some "level set function" Y_0 enjoying the following properties

$$Y_0(x,0) = 0, \quad Y_0(x,1) = 1, \quad \partial_a Y_0(a,x) > 0.$$

(As a matter of fact, this way we may recover all u_0 with a range compactly supported in $]0,1[$, and, therefore all u_0 in $L^\infty(\mathbb{T}^d)$, up to a trivial rescaling of the "flux function" Q.)

Theorem 1. *Let $Y_0(a,x)$ be any L^∞ function of $x \in T^d$ and $a \in [0,1]$ such that*

$$Y_0(x,0) = 0, \quad Y_0(x,1) = 1, \quad \partial_a Y_0(a,x) > 0.$$

Let, for all $y \in [0,1]$,

$$u_0(x,y) = \int_0^1 1\{Y_0(a,x) < y\}da,$$

Then, the unique Kruzhkov solution to the scalar conservation law

$$\partial_t u + \nabla \cdot (Q(u)) = 0,$$

with initial condition $u_0(x,y)$ can be written

$$u(t,x) = \int_0^1 1\{Y(t,a,x) < y\}da,$$

where Y solves the subdifferential inclusion in $L^2(T^d \times [0,1])$:

$$0 \ \in \ \partial_t Y + q(a) \cdot \nabla_x Y + \partial K[Y],$$

with $q = Q'$, $K[Y] = 0$ if $\partial_a Y \geq 0$, and $K[Y] = +\infty$ otherwise.

Let us be more explicit for the definition of this subdifferential inclusion.

Definition 34. *We say that Y is a solution to*

$$0 \in \partial_t Y + q(a) \cdot \nabla_x Y + \partial K[Y], \quad \text{if}:$$

1) $t \to Y(t,\cdot,\cdot) \in L^2(\mathbb{T}^d \times [0,1])$ is continuous and satisfies $\partial_a Y \geq 0$,
2) Y satisfies, in the sense of distribution,

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^d \times [0,1]} |Y - Z|^2(t,a,x)dadx$$

$$+ \int_{\mathbb{T}^d \times [0,1]} (Y - Z)(t,a,x)(\partial_t Z + q(a) \cdot \nabla_x Z)(t,a,x)dadx \leq 0,$$

for each smooth function $Z(t,a,x)$ such that $\partial_a Z \geq 0$.

Remark

As shown by Perepelitsa in [204], $Y \to F'(a) \cdot \nabla_x Y + \partial\Phi[Y]$ actually is a maximal monotone in the classical sense of [83] and generates a semi-group of contractions in L^2 . It is rather astonishing that scalar conservation laws can be reduced to the rather conventional theory of maximal monotone operators in L^2. Indeed, in the 80s, scalar conservation laws were frequently presented as one of the most striking applications of the more advanced theory or maximal operators...in L^1!

Idea of the proof

We follow the presentation of [58] rather than the earlier work of Panov [203]. (We refer to [204] for a more detailed comparison of [203] and [58].)
The main idea is to consider, instead of a single initial condition $u_0(x)$ for the scalar conservation law

$$\partial_t u + \nabla \cdot (Q(u)) = 0,$$

a one-parameter family of initial conditions $u_0(x, y)$. We make the crucial assumption that this family is monotonically increasing with respect to the parameter y, By the standard comparison principle for scalar conservation laws, the corresponding Kruzhkov solutions $u(t, x, y)$ are also monotone with respect to y. Assume, for a while, that $u(t, x, y)$ is *a priori* smooth and strictly increasing in y. Thus, we can write

$$u(t, x, Y(t, a, x)) = a, \quad Y(t, u(t, x, y)) = y$$

where $Y(t, a, x)$ is smooth and strictly increasing in $a \in [0, 1]$. Then, a straightforward calculation shows that Y must solve the simple linear equation

$$\partial_t Y + q(a) \cdot \nabla_x Y = 0$$

(which admits $Y(t, a, x) = Y(t = 0, x - tq(a), a)$ as exact solution). This is just a rephrasing of the celebrated "method of characteristics". Unfortunately, this linear equation is not able to preserve the monotonicity condition $\partial_a Y \geq 0$ in the large. However, by properly correcting it, namely by adding the subdifferential term ∂K, it is possible to enforce $\partial_a Y \geq 0$, and, this way, to recover the correct Kruzhkov entropy solutions. More precisely, as Y solves the subdifferential inclusion stated above, then

$$u(t, x, y) = \int_0^1 1\{Y(t, a, x) < y\} da$$

will be shown to be, for each fixed value y, the right entropy solution with initial conditions $x \to u_0(x, y)$.
Observe that this approach is strongly related to both the kinetic formulation and the level set method for scalar conservation laws. Let us recall that the kinetic approach amounts to lift a non-linear scalar conservation law by averaging out a linear advection equation involving a hidden extra variable. This idea

(that has obvious roots in the kinetic theory of Maxwell and Boltzmann) was introduced for scalar conservation laws in parallel by Giga-Miyakawa and the author [36, 38, 39, 148]. Its time continuous counter-part is nothing but the celebrated "kinetic formulation" of Lions, Perthame and Tadmor [178] which, with the crucial help of the so-called "averaging lemma" [151], provided the first regularity results (in suitable fractional Sobolev spaces) for multidimensional scalar conservation laws, (under suitable nonlinearity conditions). Concerning the "level set method", its application to scalar conservation laws by Tsai, Giga and Osher [228] can be interpreted as a parabolic approximation of our subdifferential inclusion, as will be discussed below.

6.3 Elements of a proof

We follow the constructive proof of [58] based on the analysis of the time-discrete scheme known as the "transport-collapse method" [39]. We will show that, as the time step goes to zero, the approximate solutions we are going to construct both converge to solutions in the Kruzhkov sense and solutions in the subdifferential sense. We assume that $Y_0(a,x) \in [0,1]$ (which is consistent with the statement of Theorem 1. We fix a time step $h > 0$ and approximate $Y(nh, a, x)$ by $Y_n(a,x)$, for each positive integer n. To get Y_n from Y_{n-1}, we perform two steps, making the following induction assumptions:

$$\partial_a Y_{n-1} \geq 0, \quad Y_{n-1} \in [0,1],$$

which are consistent with our assumptions on Y_0.

Predictor step

The first "predictor" step amounts to solve the linear equation

$$\partial_t Y + q(a) \cdot \nabla_x Y = 0,$$

for $nh - h < t < nh$, with Y_{n-1} as initial condition at $t = nh - h$. We exactly get at time $t = nh$ the predicted value:

$$Y_n^*(a,x) = Y_{n-1}(a, x - h\, q(a))$$

Thanks to the induction assumption, we still have $Y_n^* \in [0,1]$, however, although $\partial_a Y_{n-1}$ is nonnegative, the same may not be true for $\partial_a Y_n^*$. This is why, we need a "corrector step".

Corrector step

In the second step, we 'rearrange' Y^* in increasing order with respect to $a \in [0,1]$, for each fixed x, and get the corrected function Y_n. Let us recall some elementary facts about rearrangements:

Lemma 35. *Let: $a \in [0,1] \to X(a) \in \mathbb{R}$ an L^∞ function. Then, there is unique L^∞ function $Y : [0,1] \to \mathbb{R}$, such that $Y' \geq 0$ and:*

$$\int_0^1 H(y - Y(a))da = \int_0^1 H(y - X(a))da, \quad \forall y \in \mathbb{R}.$$

We say that Y is the rearrangement of X. In addition, for all $Z \in L^\infty$ such that $Z' \geq 0$, the following rearrangement inequality:

$$\int_0^1 |Y(a) - Z(a)|^p da \leq \int_0^1 |X(a) - Z(a)|^p da.$$

holds true for all $p \geq 1$.

So, we define $Y_n(a,x)$ to be, for each fixed x, the rearrangement of $Y_n^*(a,x)$ in $a \in [0,1]$:

$$\partial_a Y_n \geq 0, \quad \int_0^1 H(y - Y_n(a,x))da = \int_0^1 H(y - Y_n^*(a,x))da, \quad \forall y \in \mathbb{R}.$$

Equivalently, we may define the auxiliary function:

$$u_n(x,y) = \int_0^1 H(y - Y_n^*(a,x))da, \quad \forall y \in \mathbb{R},$$

i.e.

$$u_n(x,y) = \int_0^1 H(y - h\, Y_{n-1}(a, x - h\, q(a)))da,$$

and set:

$$Y_n(a,x) = \int_0^\infty H(a - u_n(x,y))dy.$$

At this point, Y_n is entirely determined by Y_{n-1}. Notice that, from the very definition of the rearrangement step, u_n, by definition, can be equivalently written:

$$u_n(x,y) = \int_0^1 H(y - Y_n(a,x))da.$$

Also notice that, for all function $Z(a,x)$ such that $\partial_a Z \geq 0$, and all $p \geq 1$:

$$\int |Y_n(a,x) - Z(a,x)|^p dadx \leq \int |Y_n^*(a,x) - Z(a,x)|^p dadx$$

follows from the rearrangement inequality. Finally, we see that $\partial_a Y_n \geq 0$ is automatically satisfied (this was the purpose of the rearrangement step) as well as $Y_n \in [0,1]$ (since the convex hull of the range of Y_n^* has been preserved by the rearrangement step). So, the induction assumption is enforced at step n and the scheme is well defined.

Remark

Observe that, for any fixed x, $u_n(x, y)$, as a function of y, is the (generalized) inverse of $Y_n(a, x)$, viewed as a function of a, in the sense of Lemma 35. Also notice that the level sets $\{(a, y); \quad y \geq Y_n(a, x)\}$ and $\{(a, y); \quad a \leq u_n(x, y)\}$ coincide.

6.3.1 The transport-collapse scheme revisited

The time-discrete scheme can be entirely recast in terms of the auxiliary function u_n defined as above. Indeed, introducing

$$ju_n(x, y, a) = H(u_n(x, y) - a),$$

we can rewrite the "predictor-corrector" steps in terms of u_n and ju_n as simply as:

$$u_n(x, y) = \int_0^1 ju_{n-1}(x - h\, q(a), y, a)da,$$

which exactly define the "transport-collapse" (TC) approximation to the scalar conservation law, or, equivalently, its "kinetic" approximation, according to [36, 38, 39, 148].

6.3.2 Convergence to the Kruzhkov solution

We are now going to prove that, on one hand, $Y_n(a, x)$ converges to $Y(t, a, x)$ as $nh \to t$, and, on the other hand, $u_n(x, y)$ converges to $u(t, x, y)$, where Y and u are respectively the unique solution to the subdifferential inclusion

$$0 \in \partial_t Y + q(a) \cdot \nabla_x Y + \partial K[Y],$$

with initial condition $Y_0(a, x)$ and the unique Kruzhkov solution to the scalar conservaton law with initial condition (where y is just a parameter)

$$u_0(x, y) = \int_0^1 H(y - Y_0(a, x))da. \tag{32}$$

We take for granted the convergence analysis of the TC method [36, 38, 39, 148] and obtain that, as $nh \to t$,

$$\int |u_n(x, y) - u(t, x, y)|dydx \to 0,$$

where u is the unique Kruzhkov solution with initial value u_0. More precisely, if we extend the time discrete approximations $u_n(x, y)$ to all $t \in [0, T]$ by linear interpolation in time:

$$u^h(t, x, y) = u_{n+1}(x, y)\frac{t - nh}{h} + u_n(x, y)\frac{nh + h - t}{h},$$

then $u^h - u$ converges to 0 in the space $C^0([0,T], L^1(\mathbb{R}\times\mathbb{T}^d))$ as $h \to 0$. It is now natural to introduce the level-set function Y defined from the Kruzhkov solution by

$$Y(t,a,x) = \int_0^\infty H(a - u(t,x,y))dy.$$

(Notice that, at this point, we do not know that Y is a solution to the subdifferential inclusion.) Let us interpolate the Y_n by

$$Y^h(t,a,x) = Y_{n+1}(a,x)\frac{t-nh}{h} + Y_n(a,x)\frac{nh+h-t}{h},$$

for all $t \in [nh, nh+h]$ and $n \geq 0$. Next, we crucially use the "co-area formula" (or in other words Lebesgue's "horizontal" integration by level sets) to get

$$\int |Y(t,a,x) - Y_n(a,x)|dadx = \int |u(t,x,y) - u_n(x,y)|dydx.$$

Thus:

$$\sup_{t\in[0,T]} ||Y(t,\cdot) - Y^h(t,\cdot)||_{L^1} \leq \sup_{t\in[0,T]} ||u(t,\cdot) - u^h(t,\cdot)||_{L^1} \to 0,$$

and we conclude that the approximate solution Y^h must converge to Y in $C^0([0,T], L^1([0,1]\times\mathbb{T}^d))$ as $h \to 0$. Notice that, since the Y^h are uniformly bounded in L^∞, the convergence also holds true in $C^0([0,T], L^2([0,1]\times\mathbb{T}^d))$.

We are finally left with proving that Y is the solution to the subdifferential inclusion with initial condition Y_0 in the sense of Definition 34.

6.4 Consistency of the transport-collapse scheme

Let us check that the TC scheme is consistent with the subdifferential formulation in the precise sense of Definition 34. For each smooth function $Z(t,a,x)$ with $\partial_a Z \geq 0$ and $p \geq 1$, we have

$$\int |Y_{n+1}(a,x) - Z(nh+h,a,x)|^p dadx$$

$$\leq \int |Y^*_{n+1}(a,x) - Z(nh+h,a,x)|^p dadx$$

(because of the rearrangement step, which is non expansive in any L^p)

$$= \int |Y_n(a, x - h\, q(a)) - Z(nh+h,a,x)|^p dadx$$

(by definition of the predictor step)

$$= \int |Y_n(a,x) - Z(nh+h,a,x+h\, q(a))|^p dadx$$

$$= \int |Y_n - Z(nh,\cdot)|^p dadx + h\,\Gamma + o(h)$$

where:

$$\Gamma = p\int (Y_n - Z(nh,\cdot))|Y_n - Z(nh,\cdot)|^{p-2}\{-\partial_t Z(nh,\cdot) - q\cdot\nabla_x Z(nh,\cdot)\}dadx$$

(by Taylor expanding Z about (nh,a,x)). Since the approximate solution provided by the TC scheme has a unique limit Y, as shown in the previous section, this limit must satisfy:

$$\frac{d}{dt}\int |Y-Z|^p dadx \le p\int (Y-Z)|Y-Z|^{p-2}(-\partial_t Z - q(a)\cdot\nabla_x Z)dadx,$$

in the distributional sense in t. In particular, for $p=2$, we exactly recover the differential inequality of Definition 34. We conclude that the approximate solutions generated by the TCM scheme do converge to the solutions of the subdifferential inclusion in the sense of Definition 34, which completes the proof of Theorem 1.

6.5 Viscous approximations

A natural regularization for our subdifferential inclusion amounts to substituting a barrier function for the convex cone K in $L^2([0,1]\times\mathbb{T}^d)$ of all functions Y such that $\partial_a Y \ge 0$. Typically, we introduce a convex function $\phi : \mathbb{R} \to]-\infty,+\infty]$ such that $\phi(\tau) = +\infty$ if $\tau < 0$, we define, for all $Y \in K$,

$$\Phi[Y] = \int \phi(\partial_a Y)dadx,$$

and set $\Phi[Y] = +\infty$ if Y does not belong to K. Typical examples are:

$$\phi(\tau) = -\log(\tau), \quad \phi(\tau) = \tau\log(\tau), \quad \phi(\tau) = \frac{1}{\tau}, \quad \forall\tau > 0.$$

Then, we considered the perturbed subdifferential inclusion

$$0 \ \in \ \partial_t Y + q(a)\cdot\nabla_x Y - q_0(a) + \varepsilon\partial\Phi[Y],$$

for $\varepsilon > 0$. The general theory of maximal monotone operators guarantees the convergence of the corresponding solutions to those of (**??**) as $\varepsilon \to 0$. It is not difficult (at least formally) to identify the corresponding perturbation to our scalar conservation

$$\partial_t u + \nabla\cdot(Q(u) = 0.$$

Indeed, assuming $\phi(\tau)$ to be smooth for $\tau > 0$, we get, for each smooth function Y such that $\partial_a Y > 0$:

$$\partial\Phi(Y) = -\partial_a(\phi'(\partial_a Y)).$$

Thus, any smooth solution Y of the perturbed subdifferential inclusion satisfying $\partial_a Y > 0$, solves the following parabolic equation:

$$\partial_t Y + q(a) \cdot \nabla_x Y = \varepsilon \partial_a(\phi'(\partial_a Y)).$$

Introducing, the function $u(t,x,y)$ implicitely defined by

$$u(t, Y(t,a,x), x) = a,$$

we get (by differentiating with respect to a, t and x):

$$(\partial_y u)(t, Y(t,a,x), x)\partial_a Y(t,a,x) = 1,$$

$$(\partial_t u)(t,x,y) + (\partial_y u)(t,x,y)\partial_t Y = 0,$$

$$(\nabla_x u)(t,x,y) + (\partial_y u)(t,x,y)\nabla_x Y = 0.$$

Then, we get

$$-\partial_t u - q(u) \cdot \nabla_x u - q_0(u)\partial_y u = \varepsilon \partial_y(\phi'(\frac{1}{\partial_y u})).$$

In particular, in the case $\phi(\tau) = -\log \tau$, we obtain

$$\partial_t u + q(u) \cdot \nabla_x u = \varepsilon \partial^2_{yy} u,$$

with viscosity only in the y variable. This includes viscous effects not on the space variable x but rather on the "level-set parameter" $y \in \mathbb{R}$. This unusual type of regularization has already been used and analyzed in the level-set framework developped by Giga, Giga, Osher and Tsai for scalar conservation laws [147, 228].

6.6 Related equations

A similar method can be applied to some special systems of conservation laws. A typical example (which was crucial for our understanding) is the 'Born-Infeld-Chaplygin' system considered in [52], and the related concept of 'order-preserving strings'. This system reads:

$$\partial_t(hv) + \partial_y(hv^2 - hb^2) - \partial_x(hb) = 0,$$

$$\partial_t h + \partial_y(hv) = 0, \quad \partial_t(hb) - \partial_x(hv) = 0,$$

where h, b, v are real valued functions of time t and two space variables x, y. In [52] this system is related to the following subdifferential system:

$$0 \in \partial_t Y - \partial_x W + \partial K[Y], \quad \partial_t W = \partial_x Y,$$

where (Y, W) are real valued functions of (t, a, x) and $K[Y]$ is still O or $+\infty$ according to whether $\partial_a Y \geq 0$ is true or not. The (formal) correspondence between is obtained by setting:

$$h(t, x, Y(t,a,x))\partial_a Y(t,a,x) = 1,$$

$$v(t,x,Y(t,a,x)) = \partial_t Y(t,a,x), \quad b(t,x,Y(t,a,x)) = \partial_x Y(t,a,x).$$

Unfortunately, this system is very special (its smooth solutions are easily integrable). In our opinion, it is very unlikely that L^2 formulations can be found for general hyperbolic conservation laws as easily as in the multidimensional scalar case.

6.7 More details on the subdifferential inclusion

To conclude this section, let us examine few additional properties of the subdifferential inclusion

$$0 \in \partial_t Y + q(a) \cdot \nabla_x Y + \partial K[Y],$$

obtained from the "transport-collapse" approximation scheme. First, we observe that, in the TC scheme,
1) the predictor step (a translation in the x variable by $h\ q(a)$ is isometric in all L^p spaces,
2) the corrector step (an increasing rearrangement in the a variable) is non-expansive in all L^p.
Thus the scheme is non-expansive in all $L^p([0,1] \times \mathbb{T}^d)$
Since the scheme is also invariant under translations in the x variable, we get the following a priori estimate:

$$||\nabla_x Y_n||_{L^p} \leq ||\nabla_x Y_0||_{L^p}.$$

Moreover, if we compare two solutions of the scheme Y_n and $\tilde{Y}_n = Y_{n+1}$ obtained with initial condition $\tilde{Y}_0 = Y_1$, we deduce:

$$\int |Y_{n+1}(a,x) - Y_n(a,x)|^p dadx \leq \int |Y_1(a,x) - Y_0(a,x)|^p dadx$$

$$\leq \int |Y_1^*(a,x) - Y_0(a,x)|^p dadx = \int |Y_0(a, x - h\ q(a)) - Y_0(a,x)|^p dadx.$$

So we get a second a priori estimate:

$$||Y_{n+1} - Y_n||_{L^p} \leq ||q||_{L^\infty} ||\nabla_x Y_0||_{L^p})h.$$

We conclude that the solutions Y to the subdifferential inclusion obtained from the TC scheme satisfy the a priori bounds:

$$||\nabla_x Y(t,\cdot)||_{L^p} \leq ||\nabla_x Y_0||_{L^p},$$

$$||\partial_t Y(t,\cdot)||_{L^p} \leq ||q_0||_{L^p} + ||q||_{L^\infty} ||\nabla_x Y_0||_{L^p}.$$

6.7.1 L^p and Monge-Kantorovich stability properties

As just mentioned, the solutions of the subdifferential inclusion enjoy the L^p stability property with respect to initial conditions (**??**), not only for $p = 2$ but also for all $p \geq 1$. The case $p = 1$ is of particular interest. Indeed, let us consider two solutions Y and $\tilde{Y}$ of of the subdifferential inclusion and the corresponding Kruzhkov solutions u and $\tilde{u}$, as in the proof of Theorem 1. Using the co-area formula we find, for all $t \geq 0$,

$$\int_{\mathbb{R}} \int_{\mathbb{T}^d} |u(t,x,y) - \tilde{u}(t,x,y)| dxdy =$$

$$= \int_0^1 \int_{\mathbb{R}} \int_{\mathbb{T}^d} |H(u(t,x,y) - a) - H(\tilde{u}(t,x,y) - a)| dadxdy$$

$$= \int_0^1 \int_{\mathbb{R}} \int_{\mathbb{T}^d} |H(y - Y(t,a,x)) - H(y - \tilde{Y}(t,a,x))| dadxdy$$

$$= \int_0^1 \int_{\mathbb{T}^d} |Y(t,a,x) - \tilde{Y}(t,a,x)| dxda \leq \int_0^1 \int_{\mathbb{T}^d} |Y_0(a,x) - \tilde{Y}_0(a,x)| dxda$$

$$= \int_{\mathbb{R}} \int_{\mathbb{T}^d} |u_0(x,y) - \tilde{u}_0(x,y)| dxdy.$$

Thus, Kruzhkov's L^1 stability property is nothing but a *very* incomplete output of the much stronger L^p stability property enjoyed by the subdifferential inclusion!

As a matter of fact, it is possible to translate the L^p stability of the level set function Y in terms of the Kruzhkov solution u by using Monge-Kantorovich (MK) distances. Let us first recall that for two probability measures μ and ν compactly supported on $\mathbb{R}^D$, their p MK distance can be defined (see [230] for instance), for $p \geq 1$, by:

$$\delta_p^p(\mu,\nu) = \sup \int \phi(x) d\mu(x) + \int \psi(y) d\nu(y),$$

where the supremum is taken over all pair of continuous functions ϕ and ψ such that:

$$\phi(x) + \psi(y) \leq |x - y|^p, \quad \forall x, y \in \mathbb{R}^D.$$

In dimension $D = 1$, this definition reduces to:

$$\delta_p(\mu,\nu) = ||Y - Z||_{L^p},$$

where Y and Z are respectively the "generalized inverse" of u and v defined on $\mathbb{R}$ by:

$$u(y) = \mu([-\infty, y]), \quad v(y) = \nu([-\infty, y]), \quad \forall y \in \mathbb{R}.$$

Next, observe that, for each $x \in \mathbb{T}^d$, the y derivative of the Kruzhkov solution $u(t,x,y)$, as described in Theorem **??**, can be seen as a probability measure compactly supported on $\mathbb{R}$. (Indeed, $\partial_y u \geq 0$, $u = 0$ near $y = -\infty$ and $u = 1$ near $y = +\infty$.) Then, the L^p stability property simply reads:

$$\int_{\mathbb{T}^d} \delta_p^p(\partial_y u(t,\cdot,x), \partial_y \tilde{u}(t,\cdot,x))dx \leq \int_{\mathbb{T}^d} \delta_p^p(\partial_y u_0(\cdot,x), \partial_y \tilde{u}_0(\cdot,x))dx.$$

Let us refer to [33] and [90] for recent occurences of MK distances in the field of scalar conservation laws.

Uniqueness theory

Let us consider a solution Y to the subdifferential inclusion in the sense of Definition 34. By definition $Y(t,\cdot)$ depends continuously of $t \in [0,T]$ in L^2. >From definition (34), using $Z = 0$ as a test function, we see that:

$$\frac{d}{dt}||Y(t,\cdot)||_{L^2}^2 \leq 2\int Y(t,a,x)q_0(a)\ dadx \leq ||Y(t,\cdot)||_{L^2}^2 + ||q||_{L^2}^2,$$

which implies that the L^2 norm $Y(t,\cdot)$ stays uniformly bounded on any finite interval $[0,T]$. Thus, $T > 0$ being fixed, we can mollify Y and get, for each $\epsilon \in]0,1]$ a smooth function $Y_\epsilon(t,a,x)$, still increasing in a, so that:

$$\sup_{t\in[0,T]} ||Y(t,\cdot) - Y_\epsilon(t,\cdot)||_{L^2} \leq \epsilon.$$

Let us now consider an initial condition Z_0 such that $\nabla_x Z_0$ belongs to L^2. We know that there exist a solution Z to the subdifferential inclusion, still in the sense of Definition 34. obtained by TC approximation, for which both $\partial_t Z(t,\cdot)$ and $\nabla_x Z(t,\cdot)$ stay uniformly bounded in L^2 for all $t \in [0,T]$. This function Z has enough regularity to be used as a test function when expressing that Y is a solution in the sense of Definition 34. So, for each smooth nonnegative function $\theta(t)$, compactly supported in $]0,T[$, we get from Definition 34

$$\int \{\theta'(t)|Y - Z|^2 + 2\theta(t)(Y - Z)(q_0(a) - \partial_t Z - q(a)\cdot\nabla_x Z)\}dadxdt \geq 0.$$

Substituting Y_ϵ for Y, we get

$$\int \{\theta'(t)|Y_\epsilon - Z|^2 + 2\theta(t)(Y_\epsilon - Z)(q_0(a) - \partial_t Z - q(a)\cdot\nabla_x Z)\}dadxdt \geq -C\epsilon,$$

where C is a constant depending on θ, Z, q_0 and q only. Since Z is also a solution, using Y_ϵ as a test function, we get from Definition 34:

$$\int \{\theta'(t)|Z - Y_\epsilon|^2 + 2\theta(t)(Z - Y_\epsilon)(q_0(a) - \partial_t Y_\epsilon - q(a)\cdot\nabla_x Y_\epsilon)\}dadxdt \geq 0.$$

Adding up these two inequalities, we deduce:

$$\int \{2\theta'(t)|Y_\epsilon - Z|^2 + 2\theta(t)(Y_\epsilon - Z)(\partial_t(Y_\epsilon - Z) + q(a)\cdot\nabla_x(Y_\epsilon - Z))\}dadxdt \geq -C\epsilon.$$

Integrating by part in $t \in [0, T]$ and $x \in \mathbb{T}^d$, we simply get:

$$\int \theta'(t)|Y_\epsilon - Z|^2 dadxdt \geq -C\epsilon.$$

Letting $\epsilon \to 0$, we deduce:

$$\frac{d}{dt}\int |Y - Z|^2 dadx \leq 0.$$

We conclude, at this point, that:

$$||Y(t,\cdot) - Z(t,\cdot)||_{L^2} \leq ||Y_0 - Z_0||_{L^2}, \quad \forall t \in [0, T]$$

This immediately implies the uniqueness of Y. Indeed, any other solution $\tilde{Y}$ with initial condition Y_0 must also satisfy:

$$||\tilde{Y}(t,\cdot) - Z(t,\cdot)||_{L^2} \leq ||Y_0 - Z_0||_{L^2}.$$

Thus, by the triangle inequality:

$$||\tilde{Y}(t,\cdot) - Y(t,\cdot)||_{L^2} \leq 2||Y_0 - Z_0||_{L^2}.$$

Since Z_0 is any function such that $\nabla_x Z_0$ belongs to L^2, we can make $||Y_0 - Z_0||_{L^2}$ arbitrarily small and conclude that $\tilde{Y} = Y$, which completes the proof of uniqueness.

Figure 1: time 0.0

Figure 2: time 0.6

Figure 3: time 1.2

7 Systèmes de lois conservation entropiques

Nous avons vu qu'une reformulation possible (mais pas forcément la plus pertinente) de l'équation d'Hamilton-Jacobi

$$\partial_t \phi + \frac{1}{2}|\nabla\phi|^2 = 0$$

était donnée par un système de "lois de conservation" du premier ordre, vérifié par la champ de vecteur $B = \nabla\phi$, à savoir

$$\partial_t B + \nabla(\frac{|B|^2}{2}) = 0, \quad B = B(t,x) \in \mathbb{R}^d, \quad t \geq 0, \quad x \in \mathbb{R}^d,$$

qui donne la fameuse équation de Burgers dans le cas particulier $d = 1$

$$\partial_t B + \partial_x(B^2/2) = 0, \quad B = B(t,x) \in \mathbb{R}, \quad t \geq 0, \quad x \in \mathbb{R}.$$

De façon plus générale, on appelle "système de lois de conservation du premier ordre tout système d'EDP de la forme

$$\partial_t U^\alpha + \partial_i(\mathcal{F}^{i\alpha}(U)) = 0, \;\; \alpha = 1, \cdots, m,$$

(avec sommation implicite sur les indices répétés) où $U = U(t,x) \in \mathcal{W} \subset \mathbb{R}^m$, $t \geq 0$, $x \in \mathbb{R}^d$, $\partial_t = \frac{\partial}{\partial t}$, $\partial_i = \frac{\partial}{\partial x_i}$, $\mathcal{W}$ est un ouvert convexe lisse, et où la "fonction de flux" (c'est le nom consacré) $\mathcal{F} : \mathcal{W} \to \mathbb{R}^{d\times m}$ est lisse avec certaines propriétés de croissance contrôlée au bord de $\mathcal{W}$. Encore une fois, on peut remonter à Euler pour fonder la théorie, avec son équation de la dynamique des gaz, qui s'écrit, dans le cas "isotherme",

$$\partial_t \rho + \nabla \cdot q = 0, \;\; \partial_t q + \nabla \cdot (\frac{q \otimes q}{\rho}) + \nabla\rho = 0$$

($\rho > 0$ et $q \in \mathbb{R}^d$ notant respectivement les champs de densité et de quantité de mouvement du gaz). On a alors

$$U = (\rho, q) \in \mathcal{W} =]0, +\infty[\times\mathbb{R}^d, \quad \mathcal{F}(U) = (q, \frac{q \otimes q}{\rho} + \rho\, I_d).$$

On va dorénavant se limiter à la classe restreinte des "systèmes de lois de conservation avec entropie convexe", ou, plus brièvement "systèmes entropiques de lois de conservation" (SELC).

Definition 36. *On appelle SELC un système de lois de conservation pour lequel la fonction de flux $\mathcal{F}$ vérifie la condition de symétrie supplémentaire*

$$\forall i \in \{1, \cdots, n\}, \;\; \forall \beta, \gamma \in \{1, \cdots, m\}, \;\; \partial^2_{\alpha\beta}\mathcal{E}\partial_\gamma\mathcal{F}^{i\alpha} = \partial^2_{\alpha\gamma}\mathcal{E}\partial_\beta\mathcal{F}^{i\alpha},$$

pour une certaine fonction lisse, appelée "entropie", $\mathcal{E} : \mathcal{W} \to \mathbb{R}$, strictement convexe au sens que $(\partial^2_{\alpha\beta}\mathcal{E})$ est une matrice définie positive en tout point de $\mathcal{W}$.

Cette propriété (a priori étrange) entraîne la "conservation de l'entropie" au sens que toute solution U, de classe C^1, du système de lois de conservation est solution de la loi de conservation supplémentaire

$$\partial_t(\mathcal{E}(U)) + \partial_i(\mathcal{Q}^i(U)) = 0,$$

où la fonction "de flux d'entropie" $\mathcal{Q} : \mathcal{W} \to \mathbb{R}^d$ peut se déduire de $\mathcal{F}$ et $\mathcal{E}$.

[En effet, la condition de symétrie relativement à $\mathcal{E}$ équivaut à

$$\partial_\gamma(\partial_\alpha\mathcal{E}\partial_\beta\mathcal{F}^{i\alpha}) = \partial_\beta(\partial_\alpha\mathcal{E}\partial_\gamma\mathcal{F}^{i\alpha})$$

qui signifie que $\partial_\alpha\mathcal{E}\partial_\beta\mathcal{F}^{i\alpha}$ est le gradient d'une certaine fonction $Q^i : \mathcal{W} \to \mathbb{R}$, i.e. $\partial_\alpha\mathcal{E}\partial_\beta\mathcal{F}^{i\alpha} = \partial_\beta Q^i$. Ainsi, toute solution C^1 U satisfait

$$-\partial_t(\mathcal{E}(U)) = \partial_\alpha\mathcal{E}(U)\partial_i(\mathcal{F}^{i\alpha}(U)) = \partial_\alpha\mathcal{E}(U)\partial_\beta\mathcal{F}^{i\alpha}(U)\partial_i U^\beta = \partial_\beta\mathcal{Q}^i(U)\partial_i U^\beta = \partial_i(\mathcal{Q}^i(U)),$$

ce qui est exactement la loi de conservation de l'entropie.]

Cette classe d'EDP contient de nombreux exemples issus de la mécanique des milieux continus, de la physique et de la géométrie (équations d'Euler des fluides compressibles, élastodynamique, magnétohydrodynamique, électromagnétisme, surfaces extrémales de l'espace de Minkowski, cordes classiques etc...). Comme déjà mentionné, l'exemple le plus simple est l'équation de Burgers (sans viscoisté)

$$\partial_t u + \partial_x(\frac{u^2}{2}) = 0, \quad u \in \mathbb{R}, \tag{33}$$

où on a $\mathcal{F}(u) = u^2/2$ et pour laquelle on peut poser $\mathcal{E}(u) = u^2/2$, $\mathcal{Q}(u) = u^3/3$.

Plus générale est la classe des "lois de conservation scalaires", pour lesquelles $m = 1$ avec $\mathcal{W} = \mathbb{R}$, et, du coup, la condition de symétrie est trivialement satisfaite, pour toute fonction convexe $\mathcal{E}$, comme on pourra facilement le vérifier. Cette classe a des propriétés similaire à celle des équations d'Hamilton-Jacobi et pourra faire l'objet d'une étude plus détaillée car on en a une théorie très complète (existence de solutions globales, en un sens à préciser -on parlera de "solutions entropiques"-, unicité et stabilité dans L^1 des solutions par rapport à leurs données initiales).

Le cas des équations d'Euler est plus riche. On pourra vérifier (non sans calcul!) que, par exemple dans le cas isotherme, on a bien une entropie strictement convexe, à savoir,

$$\mathcal{E}(U) = \frac{|q|^2}{2\rho} + \rho(\log\rho - 1), \quad U = (\rho, q).$$

7.1 Principe de moindre action et équations d'ondes

Une sous classe des SELC peut être obtenue à partir du "principe de moindre action" appliqué, non plus à des trajectoires (voire des flots) comme on l'a vu dans des cours précédents, mais à des "champs scalaires". On se donne un "lagrangien", i.e. une fonction défini sur $\mathbb{R} \times \mathbb{R}^d$,

$$(E, B) \in \mathbb{R} \times \mathbb{R}^d \to L(E, B) \in \mathbb{R}$$

et l'hypothèse la plus importante est que $L(E, B)$ est strictement convexe en $E \in \mathbb{R}$. Pour simplifier la discussion, on supposera L de la forme

$$L(E, B) = \sup_{D \in \mathbb{R}} DE - H(D, B)$$

où H appelée "hamiltonien" est lisse, avec une constante $r \in]0, 1]$ telle que

$$r \leq \partial_D^2 H(D, B) \leq 1/r, \quad \forall (D, B) \in \mathbb{R} \times \mathbb{R}^d.$$

On dira qu'une fonction ϕ ("champ scalaire") définie sur un ouvert Ω de $\mathbb{R} \times \mathbb{R}^d$ satisfait le principe de moindre action relativement à L si, pour toute "perturbation" ψ dans $C_c^\infty(\Omega)$, on a

$$\frac{d}{d\epsilon}_{|\epsilon=0} \int_\Omega (L(\partial_t \phi(t,x) + \epsilon \partial_t \psi(t,x), \nabla \phi(t,x) + \epsilon \nabla \psi(t,x)) L(\partial_t \phi(t,x), \nabla \phi(t,x)))\, dt dx = 0.$$

(On garde ici la terminologie de "moindre action", datant du 18ème siècle, bien qu'il s'agisse en fait de points critiques et non pas de minima.) Si ϕ est de classe C^2, cela se traduit par l'EDP d'ordre 2, qu'on peut voir comme équation d'ondes non-linéaires :

$$\partial_t((\partial_E L)(\partial_t \phi(t,x), \nabla \phi(t,x))) + \partial_i((\partial_{B_i} L)(\partial_t \phi(t,x), \nabla \phi(t,x))) = 0,$$

qui n'est pas très facile à manipuler telle quelle. L'hamiltonien H va nous permettre une écriture bien plus commode. Compte tenu des hypothèses, on a de façon élémentaire :

Lemma 37. *Pour chaque $B \in \mathbb{R}^d$ fixé,*

$$D \in \mathbb{R} \to \xi(D, B) = \partial_D H(D, B) \in \mathbb{R}$$

définit un difféomorphisme croissant de $\mathbb{R}$ et on a les propriétés suivantes :

$$L(\xi(D, B), B) = \xi(D, B) D - H(D, B)$$

$$\partial_E L(\xi(D, B), B) = D, \quad \partial_{B_i} L(\xi(D, B), B) = -\partial_{B_i} H(D, B).$$

Ceci nous permet de réécrire l'EDP comme un système de lois de conservation, en introduisant les champs

$$(t, x) \in \Omega \to (D(t, x), B(t, x)) \in \mathbb{R} \times \mathbb{R}^d$$

obtenus en posant

$$B_i(t, x) = \partial_i \phi(t, x), \quad D(t, x) = (\partial_E L)(\partial_t \phi(t, x), \nabla \phi(t, x)).$$

ce qui (par le lemme) donne

$$\partial_t \phi(t,x) = (\partial_D H)(D(t,x), B(t,x)).$$

On voit alors que l'équation du second ordre régissant le champ ϕ peut se réécrire comme système de lois de conservation pour (D, B) :

$$\partial_t B(t,x) = \partial_i \left((\partial_D H)(D(t,x), B(t,x)) \right), \quad \partial_t D(t,x) = \partial_i \left((\partial_B H)(D(t,x), B(t,x)) \right).$$

On obtient alors, pour toute solution classique ϕ,

$$\partial_t \left(H(D(t,x), B(t,x)) \right) = \partial_i \left((\partial_D H)(D(t,x), B(t,x))(\partial_{B_i} H)(D(t,x), B(t,x)) \right).$$

Si on suppose, de plus, que $H(D, B)$ est strictement convexe on aura finalement obtenu un SELC avec "entropie" H et "flux d'entropie"

$$\mathcal{Q}^i(D, B) = -\partial_D H(D, B) \partial_{B_i} H(D, B).$$

Evidemment, le cas le plus simple est obtenue quand

$$H(D, B) = \frac{D^2 + |B|^2}{2},$$

qui donne

$$L(E, B) = \frac{E^2 - |B|^2}{2},$$

et correspond à l'équation des ondes linéaire

$$\partial^2_{tt} \phi(t,x) = \Delta \phi(t,x)).$$

7.2 Quelques résultats sur les systèmes entropiques de lois de conservation

On va montrer un certain nombre de résultats généraux, dans le cas idéalisé, où on suppose, pour simplifier les preuves, quelques propriétés supplémentaires (qui ne sont pas strictement parlant satisfaites par les exemples mentionnés, comme les équations d'Euler et même celle de Burgers, mais tel est souvent le sort des théorèmes "généraux" !). Ainsi on supposera
i) $\mathcal{W} = \mathbb{R}^m$;
ii) Toutes les dérivées de F sont bornées;
iii) Il existe une constante $r \in]0, 1]$ telle qu'en tout point de $\mathcal{W} = \mathbb{R}^m$, le spectre de la matrice $\partial^2_{\alpha\beta} \mathcal{E}$ est contenu dans $[r, 1/r]$,
et on se limitera aux solutions $U = U(t,x)$ qui sont $\mathbb{Z}^d-$périodiques en espace (autrement dit $x \in \mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$).
Une première propriété, purement structurelle, est l'écriture du système entropique de lois conservation (SELC), sous forme symétrique :

Theorem 6. *Pour toute solution $U = U(t,x)$ du SELC, de classe C^1 sur $[0,T] \times \mathbb{T}^d$, on a le système symétrique du premier ordre,*

$$A^0_{\alpha\beta}(t,x)\partial_t U^\beta(t,x) + A^j_{\alpha\gamma}(t,x)\partial_j U^\gamma(t,x) = 0,$$

où les A^0, A^j, $j = 1, \cdots m$, sont des champs de matrices symétriques $m \times m$, définies positives dans le cas des A^0.

Cette écriture "symétrique" du SELC est importante car elle est le point de départ du résultat d'existence et d'unicité de solutions classiques en temps petit :

Theorem 7. *Pour toute donnée initiale U_0 dans $H^s(\mathbb{T}^d)$, avec $s - d/2 > 1$, il existe un temps $T > 0$ (qui dépend de U_0) tel que le SELC admette une unique solution $U = U(t,x)$ de classe C^1 en $(t,x) \in [0,T] \times \mathbb{T}^d$ de donnée initiale U_0: $U(0,\cdot) = U_0$.*

Notons que l'exposant $s-d/2 > 1$ est celui qui garantit l'injection de $H^s(\mathbb{T}^d)$ dans $C^1(\mathbb{T}^d)$. On examine ensuite le lien entre solutions classiques et solutions faibles, définies comme suit :

Definition 38. *On appelle solution faible du SELC de donnée initiale U_0 sur un intervalle de temps $[0,T]$ toute fonction $U \in L^2([0,T] \times \mathbb{T}^d; \mathbb{R}^m)$ telle que*

$$\int_{[0,T]\times\mathbb{T}^d} \partial_t W_\alpha U^\alpha + \partial_i W_\alpha \mathcal{F}^{i\alpha}(U) + \int_{\mathbb{T}^d} W_\alpha(0,\cdot)U_0^\alpha = 0,$$

pour toute fonction $(t,x) \in [0,T] \times \mathbb{T}^d \to W = W(t,x) \in \mathbb{R}^m$, de classe C^∞, telle que $W(T,\cdot) = 0$.

(Le choix de l'espace L^p avec $p = 2$ est circonstanciel et lié aux hypothèses "idéalisées". Dans les applications concrètes, l'exposant p peut changer.)

Theorem 8. *Soit U solution du SELC, de classe C^1 sur $[0,T]\times\mathbb{T}^d$ et de donnée initiale U_0. Alors U est l'unique solution faible de donnée initiale U_0, vérifiant*

$$\int_{\mathbb{T}^d} \mathcal{E}(U(t,x))dx \leq \int_{\mathbb{T}^d} \mathcal{E}(U_0(x))dx,$$

pour presque tout $t \in [0,T]$.

Dans cet énoncé, dit "d'unicité fort-faible", la condition que l'entropie de la solution faible reste presque toujours inférieure à celle de la donnée initiale joue un rôle clé.

[En effet, la méthode dite "d'intégration convexe" (héritée de Nash et Gromov et mise au point dans le contexte des équations d'Euler, puis de certains SELC, par De Lellis et Székelyhidi il y a une dizaine d'années) montre, au moins dans certains cas emblématiques, qu'on peut générer quantité de solutions faibles pour une donnée initiale fixée ! Il faut donc un critère supplémentaire pour assurer l'unicité "fort-faible".]

Preuve du Théorème 6

Soit U solution du SELC, de classe C^1 sur $[0,T] \times \mathbb{T}^d$. Comme on a

$$\partial_t(\mathcal{E}_{,\alpha}(U)) = \mathcal{E}_{,\alpha\beta}(U)\partial_t U^\beta = \mathcal{E}_{,\alpha\beta}(U)\mathcal{F}^{j\beta}_{,\gamma}(U)\partial_j U^\gamma$$

(où on a noté les dérivées partielles par rapport à $U \in \mathbb{R}^m$ par des virgules), il suffit pour prouver le théorème de poser

$$A^0_{\alpha\beta}(t,x) = \mathcal{E}_{,\alpha\beta}(U(t,x))$$

$$\begin{aligned} A^j_{\alpha\gamma}(t,x) &= \mathcal{E}_{,\alpha\beta}(U(t,x))\mathcal{F}^{j\beta}_{,\gamma}(U(t,x)) \\ &= \mathcal{E}_{,\gamma\beta}(U(t,x))\mathcal{F}^{j\beta}_{,\alpha}(U(t,x)) \end{aligned}$$

(par la condition de symétrie caractérisant les SELC, au côté de la convexité de $\mathcal{E}$).

Fin de la preuve.

Eléments de preuve du Théorème 7

Le point de départ est un résultat de stabilité dans l'espace $L^2(\mathbb{T}^d)$, et plus généralement dans les espaces de Sobolev $H^s(\mathbb{T}^d)$, du système LINEAIRE à coefficients variables

$$A^0_{\alpha\beta}(t,x)\partial_t U^\beta(t,x) + A^j_{\alpha\gamma}(t,x)\partial_j U^\gamma(t,x) = M_{\alpha\gamma}(t,x)U^\gamma(t,x)$$

où les M, A^0, A^j, $j = 1, \cdots m$, sont des champs DONNES de matrices $m \times m$, symétriques dans le cas des A^j et symétriques définies positives dans le cas des A^0. Une fois acquis ce résultat, on pourra aborder le système non-linéaire où les matrices A^k dépendent de l'inconnue U via les formules (vues plus haut)

$$A^0_{\alpha\beta}(t,x) = \mathcal{E}_{,\alpha\beta}(U(t,x))$$

$$A^j_{\alpha\gamma}(t,x) = \mathcal{E}_{,\gamma\beta}(U(t,x))\mathcal{F}^{j\beta}_{,\alpha}(U(t,x)),$$

par un argument de point fixe, en "contrôlant" les non-linéarités par la norme $C^1(\mathbb{T}^d)$ de U, qui est elle-même contrôlée par la norme $H^s(\mathbb{T}^d)$ de U, dès que (grâce aux "injections de Sobolev") $s - d/2 > 1$. La preuve complète étant assez technique, on se contentera d'établir la "stabilité L^2" du système linéaire à coefficients variables :

Proposition 39. *Supposons que $(t,x) \to A^0(t,x)$ et $(t,x) \to C(t,x)$ sont des champs donnés de matrices symétriques de classe C^1 sur $\mathbb{R} \times \mathbb{T}^d$ et qu'il existe des constantes $r \in]0,1]$ et $\kappa \geq 0$ telles que, en tout point (t,x), le spectre de $A^0(t,x)$ et celui de $C(t,x)$ sont respectivement contenus dans $[r, 1/r]$ et $[0,\kappa]$. Alors le système linéaire d'EDP du 1er ordre*

$$A^0_{\alpha\beta}(t,x)\partial_t U^\beta(t,x) + A^j_{\alpha\gamma}(t,x)\partial_j U^\gamma(t,x) = M_{\alpha\gamma}(t,x)U^\gamma(t,x)$$

génère des solutions U, stables dans l'espace $L^2(\mathbb{T}^d)$, au sens que

$$||U(t,\cdot)||_{L^2(\mathbb{T}^d)} \leq ||U(s,\cdot)||_{L^2(\mathbb{T}^d)} \exp(\kappa|t-s|)/r^2, \quad \forall t,s \in \mathbb{R}.$$

En multipliant le système par U^α et en sommant en α on a, au moins formellement,

$$\partial_t \left(U^\alpha A^0_{\alpha\beta} U^\beta\right) - \partial_j \left(U^\alpha A^j_{\alpha\beta} U^\beta\right) = U^\alpha C_{\alpha\beta} U^\beta,$$

où

$$C = \partial_t A^0 - \partial_j A^j + M + M^T,$$

En intégrant en $x \in \mathbb{T}^d$, on obtient donc

$$\frac{d}{dt} \int_{\mathbb{T}^d} U^\alpha A^0_{\alpha\beta} U^\beta = \int_{\mathbb{T}^d} U^\alpha C_{\alpha\beta} U^\beta.$$

Comme on suppose que le spectre de C est contenu dans $[0, \kappa]$ et celui de A^0 dans $[r, 1/r]$, on déduit

$$|\frac{d}{dt} \int_{\mathbb{T}^d} U^\alpha A^0_{\alpha\beta} U^\beta| \leq \kappa/r \int_{\mathbb{T}^d} U^\alpha A^0_{\alpha\beta} U^\beta$$

et donc

$$\int_{\mathbb{T}^d} U^\alpha(t,\cdot) A^0_{\alpha\beta}(t,\cdot) U^\beta(t,\cdot) \leq \exp(\kappa|t-s|/r) \int_{\mathbb{T}^d} U^\alpha(s,\cdot) A^0_{\alpha\beta}(s,\cdot) U^\beta(s,\cdot),$$

d'où finalement

$$||U(t,\cdot)||_{L^2(\mathbb{T}^d)} \leq ||U(s,\cdot)||_{L^2(\mathbb{T}^d)} \exp(\kappa|t-s|/r)/r^2, \quad \forall t, s \in \mathbb{R}.$$

N.B. On obtient (plus laborieusement) le même type d'estimation pour la norme H^s pour s entier positif et on peut, à partir de $s - d/2 > 1$ ainsi contrôler la norme C^1 de U (ce qui est crucial pour le passage au non-linéaire).

Fin de l'esquisse de preuve.

Preuve du Théorème 8

Soit $(t,x) \in [0,T] \times \mathbb{T}^d \to U(t,x) \in \mathbb{R}^m$ une solution faible du SELC au sens de la Definition 38 et soit $(t,x) \in [0,T] \times \mathbb{T}^d \to V(t,x) \in \mathbb{R}^m$ une fonction C^∞. On introduit les fonctions

$$\eta(u,v) = \mathcal{E}(u) - \mathcal{E}(v) - \mathcal{E}_{,\alpha}(v)(u^\alpha - v^\alpha) \quad \forall u, v \in \mathbb{R}^m,$$

$$\zeta^{i\alpha}(u,v) = \mathcal{F}^{i\alpha}(u) - \mathcal{F}^{i\alpha}(v) - \mathcal{F}^{i\alpha}_{,\gamma}(v)(u^\gamma - v^\gamma) \quad \forall u, v \in \mathbb{R}^m, \quad i \in \{1, \cdots, d\}, \quad \alpha \in \{1, \cdots, m\}.$$

Compte tenu des hypothèses sur $\mathcal{E}$ et $\mathcal{F}$, on a

$$r|u-v|^2 \leq \eta(u,v) \leq |u-v|^2/r, \quad |\zeta(u,v)| \leq C\eta(u,v)$$

(où C est une constante qui dépend de la norme du sup des dérivées secondes de $\mathcal{F}$) et, ainsi, la quantité

$$\int_{\mathbb{T}^d} \eta(U(t,x), V(t,x))dx$$

va nous permettre d'estimer

$$||U(t,\cdot) - V(t,\cdot)||_{L^2}^2.$$

Calculons, au sens des distributions sur $]0,T[\times\mathbb{T}^d$, $\partial_t(\eta(U,V))$, et, pour commencer,

$$\partial_t\left(\mathcal{E}(V) + \mathcal{E}_{,\alpha}(V)(U^\alpha - V^\alpha)\right)$$
$$= \mathcal{E}_{,\alpha}(V)\partial_t V^\alpha + \mathcal{E}_{,\alpha\beta}(V)\partial_t V^\beta(U^\alpha - V^\alpha) + \mathcal{E}_{,\alpha}(V)(-\partial_i(\mathcal{F}^{i\alpha}(U)) - \partial_t V^\alpha)$$

(en utilisant que U est solution du SELC au sens faible et donc au sens des distributions, ce qui justifie l'écriture du terme $\mathcal{E}_{,\alpha}(V)\partial_i(\mathcal{F}^{i\alpha}(U))$ au sens des distributions)

$$= \mathcal{E}_{,\alpha\beta}(V)(R^\beta[V] - \mathcal{F}^{i\beta}{}_{,\gamma}(V)\partial_i V^\gamma)(U^\alpha - V^\alpha)$$
$$-\partial_i(\mathcal{E}_{,\alpha}(V)\mathcal{F}^{i\alpha}(U)) + \mathcal{E}_{,\alpha\gamma}(V)\partial_i V^\gamma \mathcal{F}^{i\alpha}(U)$$

[où on a introduit le "résidu"

$$R^\beta[V] = \partial_t V^\beta + \partial_i(\mathcal{F}^{i\beta}(V)) = \partial_t V^\beta + \mathcal{F}^{i\beta}{}_{,\gamma}(V)\partial_i V^\gamma$$

qui fait de $V \to R[V]$ un opérateur non-linéaire dont on remarque qu'il s'annule dès que V est solution C^1 du SELC, ce qu'on mettra à profit plus loin]

$$= \mathcal{E}_{,\alpha\beta}(V)(U^\alpha - V^\alpha)R^\beta[V] - \mathcal{E}_{,\gamma\beta}(V)\mathcal{F}^{i\beta}{}_{,\alpha}(V)\partial_i V^\gamma(U^\alpha - V^\alpha)$$
$$-\partial_i(\mathcal{E}_{,\alpha}(V)\mathcal{F}^{i\alpha}(U)) + \mathcal{E}_{,\beta\gamma}(V)\partial_i V^\gamma \mathcal{F}^{i\beta}(U)$$

(où on a utilisé la propriété de symétrie de $\mathcal{F}$ relativement à $\mathcal{E}$ et aussi remplacé l'indice muet α par β dans le tout dernier terme)

$$= \mathcal{E}_{,\alpha\beta}(V)(U^\alpha - V^\alpha)R^\beta[V] + \mathcal{E}_{,\gamma\beta}(V)\partial_i V^\gamma(\zeta^{i\beta}(U,V) + \mathcal{F}^{i\beta}(V)) - \partial_i(\mathcal{E}_{,\alpha}(V)\mathcal{F}^{i\alpha}(U))$$

(où on a utilisé la définition de ζ). Notons que, par définition de $\mathcal{Q}$,

$$\mathcal{E}_{,\gamma\beta}(V)\partial_i V^\gamma \mathcal{F}^{i\beta}(V) = \partial_i\left(\mathcal{E}_{,\beta}(V)\mathcal{F}^{i\beta}(V)\right) - \mathcal{F}^{i\beta}_{,\gamma}(V)\mathcal{E}_{,\beta}(V)\partial_i V^\gamma$$
$$= \partial_i\left(\mathcal{E}_{,\beta}(V)\mathcal{F}^{i\beta}(V) - \mathcal{Q}^i(V)\right).$$

On a ainsi obtenu, au sens des distributions sur $]0,T[\times\mathbb{T}^d$,

$$\partial_t\left(\mathcal{E}(V) + \mathcal{E}_{,\alpha}(V)(U^\alpha - V^\alpha)\right)$$
$$= \mathcal{E}_{,\alpha\beta}(V)(U^\alpha - V^\alpha)R^\beta[V] + \mathcal{E}_{,\gamma\beta}(V)\partial_i V^\gamma \zeta^{i\beta}(U,V) - \partial_i\left(\mathcal{Q}^i(V)\right).$$

Comme U est solution faible au sens de la définition 38, on peut exprimer cette équation sous forme intégrale tout en y incorporant la donnée initiale U_0. Ce faisant, on trouve en particulier, pour toute fonction test $\psi(t,x)=\chi(t)\otimes 1$ avec $\chi\in C^\infty(\mathbb{R})$ à support dans $]-\infty,T[$,

$$-\int_0^T \chi'(t)\int_{\mathbb{T}^d}\left(\mathcal{E}(V)+\mathcal{E}_{,\alpha}(V)(U^\alpha-V^\alpha)\right)(t,x)dxdt$$

$$-\chi(0)\int_{\mathbb{T}^d}\left(\mathcal{E}(V(0,x))+\mathcal{E}_{,\alpha}(V(0,x))(U_0^\alpha(x)-V^\alpha(0,x))\right)dx$$

$$=\int_0^T \chi(t)\int_{\mathbb{T}^d}\left(\mathcal{E}_{,\alpha\beta}(V)(U^\alpha-V^\alpha)R^\beta[V]+\mathcal{E}_{,\gamma\beta}(V)\partial_i V^\gamma\zeta^{i\beta}(U,V)\right)(t,x)dxdt.$$

C'est ici qu'on va incorporer le terme $\mathcal{E}(U)$ dans le membre de gauche de façon à faire apparaitre à gauche le terme

$$\eta(U,V)=\mathcal{E}(U)-\mathcal{E}(V)-\mathcal{E}_{,\alpha}(V)(U^\alpha-V^\alpha).$$

On trouve (en changeant tous les signes)

$$-\int_0^T \chi'(t)\int_{\mathbb{T}^d}\eta(U,V)(t,x)dxdt=-\int_0^T\chi'(t)\int_{\mathbb{T}^d}\mathcal{E}(U)(t,x)dxdt$$

$$-\chi(0)\int_{\mathbb{T}^d}\eta(U_0(x),V(0,x))dx+\chi(0)\int_{\mathbb{T}^d}\mathcal{E}(U_0(x))dx$$

$$-\int_0^T\chi(t)\int_{\mathbb{T}^d}\mathcal{E}_{,\alpha\beta}(V)(U^\alpha-V^\alpha)R^\beta[V](t,x)dxdt$$

$$-\int_0^T\chi(t)\int_{\mathbb{T}^d}\mathcal{E}_{,\gamma\beta}(V)\partial_iV^\gamma\zeta^{i\beta}(U,V)(t,x)dxdt.$$

Compte tenu des hypothèses faites sur $\mathcal{E}$ et $\mathcal{F}$, on peut, en supposant dorénavant $\chi\geq 0$, facilement majorer le tout dernier terme par

$$c\int_0^T\chi(t)\lambda(t)\int_{\mathbb{T}^d}\eta(U,V)(t,x)dxdt,$$

où on note $\lambda(t)$ la constante de Lipschitz en $x\in\mathbb{T}^d$ de $V(t,\cdot)$ et par c une constante dépendant seulement des fonctions $\mathcal{E}$ et $\mathcal{F}$. En notant temporairement

$$\theta(t)=\int_{\mathbb{T}^d}\eta(U,V)(t,x)dx,\quad h(t)=\int_{\mathbb{T}^d}\mathcal{E}(U(t,x))dx,$$

$$\theta_0=\int_{\mathbb{T}^d}\eta(U_0(x),V(0,x))dx,\quad h_0=\int_{\mathbb{T}^d}\mathcal{E}(U_0(x))dx,$$

$$\rho(t)=\int_{\mathbb{T}^d}\left(\mathcal{E}_{,\alpha\beta}(V)(U^\alpha-V^\alpha)R^\beta[V]\right)(t,x)dx$$

on a donc obtenu

$$-\int_0^T \chi'(t)\theta(t)dt \leq -\int_0^T \chi'(t)h(t)dt+\chi(0)(\theta_0-h_0)-\int_0^T \chi(t)\rho(t)dt+c\int_0^T \chi(t)\lambda(t)\theta(t)dt.$$

Presque tout $\tau \in [0,T[$ est un point de Lebesgue de la fonction θ et de la fonction h. En un tel point, que l'on fixe, on prend $\epsilon > 0$ assez petit pour que $\tau + \epsilon < T$ et on prend $\chi \in C_c^\infty(\mathbb{R}^)$ de sorte que :
i) pour $t \in [-1, \tau - \epsilon]$, $\chi(t) = 1$;
ii) pour $t > \tau + \epsilon$, $\chi(t) = 0$;
iii) pour $t \in [\tau - \epsilon, \tau + \epsilon]$, $\chi(t)$ est décroissante. A la limite $\epsilon \downarrow 0$, on trouve alors

$$\theta(\tau) \leq h(\tau) + \theta_0 - h_0 - \int_0^\tau \rho(t)dt + c\int_0^\tau \lambda(t)\theta(t)dt.$$

C'est ici que nous utilisons crucialemnt l'hypothèse que

$$\int_{\mathbb{T}^d} \mathcal{E}(U)(\tau, x)dx \leq \int_{\mathbb{T}^d} \mathcal{E}(U_0(x))dx$$

pour presque tout $\tau \in [0,T]$, i.e. $h(\tau) \leq h_0$. On en déduit, pour presque tout $\tau \in [0,T[$,

$$\theta(\tau) \leq \theta_0 - \int_0^\tau \rho(t)dt + c\int_0^\tau \lambda(t)\theta(t)dt$$

et, par le lemme de Grönwall, on a donc obtenu :

Proposition 40. *Pour presque tout $t \in [0,T]$,*

$$\theta(t) \leq \theta_0 \exp(c\int_0^t \lambda(s)ds) - \int_0^t \rho(s)\exp(c\int_s^t \lambda(\sigma)d\sigma)ds.$$

où $\lambda(t)$ est la constante de Lipschitz en $x \in \mathbb{T}^d$ de $V(t,\cdot)$, c une constante dépendant seulement des fonctions $\mathcal{E}$ et $\mathcal{F}$ et

$$\theta(t) = \int_{\mathbb{T}^d} \eta(U,V)(t,x)dx, \quad \theta_0 = \int_{\mathbb{T}^d} \eta(U_0(x), V(0,x))dx,$$

$$\rho(t) = \int_{\mathbb{T}^d} \left(\mathcal{E}_{,\alpha\beta}(V)(U^\alpha - V^\alpha)R^\beta[V]\right)(t,x)dx.$$

Si on suppose que V est une solution lisse du SELC avec donnée initiale U_0, on a automatiquement $R[V] = 0$, puisque

$$R^\beta[V] = \partial_t V^\beta + \partial_i(\mathcal{F}^{i\beta}(V)),$$

et $\theta_0 = 0$. On a donc

$$\int_{\mathbb{T}^d} \eta(U,V)(t,x)dx = 0,$$

pour presque tout $t \in [0,T]$. Comme cette quantité domine, à une constante près, le carré de la norme L^2 de $U(t,\cdot) - V(t,\cdot)$, on en conclut que $U = V$ ce qui montre bien l'unicité de V parmi toutes les solutions faibles issues de U_0 et dont l'entropie au temps t ne dépasse pas celle de U_0, presque surement. Ceci termine la preuve du théorème 8.

Fin de la preuve du théorème 8

7.3 Le concept de "solutions dissipatives"

Au cours de la démonstration du théorème 8, on a établi la proposition 40 qui nous suggère un nouveau concept de solution généralisée pour le SELC. Cette idée remonte à DiPerna dans les années 1980 et, plus explicitement, à Lions dans les années 1990 (dans le cas des équations d'Euler, en régime incompressible, ce qui sort du cadre strict des SELC -mais en est en fait juste un cas limite-). On remarque que l'inégalité énoncée dans la proposition, est *convexe* relativement à U. En effet $\eta(U,V)$ est convexe en U par définition et au second membre ne figurent que des termes linéaires en U. C'est une propriété remarquable qui fournit facilement de la "compacité faible". Plus précisément, considérons l'espace, qu'on note $C^0_w([0,T], L^2(\mathbb{T}^d;\mathbb{R}^m)$ de toutes les fonctions

$$U : t \in [0,T] \to U(t,\cdot) \in L^2(\mathbb{T}^d;\mathbb{R}^m)$$

qui sont continues par rapport à t relativement à la topologie faible de $L^2(\mathbb{T}^d;\mathbb{R}^m)$, i.e. telles que pour toute fonction $\psi \in L^2(\mathbb{T}^d;\mathbb{R}^m)$,

$$t \in [0,T] \to \int_{\mathbb{T}^d} U^\alpha(t,x)\psi_\alpha(x)dx$$

est une fonction continue.

Definition 41. *On dit qu'une fonction $U \in C^0_w([0,T], L^2(\mathbb{T}^d;\mathbb{R}^m))$ est une solution "dissipative" (au sens de DiPerna et Lions) du SELC avec donnée initiale U_0 si elle vérifie $U(0,\cdot) = U_0$ et l'inégalité de la proposition 40 pour toute fonction lisse V.*

On a alors

Proposition 42. *Etant donné $U_0 \in L^2(\mathbb{T}^d;\mathbb{R}^m)$, considérons l'ensemble des solutions dissipatives du SELC avec donnée initiale U_0. Alors :*
i) dans la mesure où il n'est pas vide, il est convexe;
ii) dès que le SELC admet une solution lisse de donnée initiale U_0, l'ensemble est un singleton et se réduit à cette solution.

Ce résultat est loin d'être satisfaisant. Néanmoins, on a deux constats intéressants :
i) il est souvent assez facile (bien que parfois passablement technique), par un procédé d'approximation bien choisi, de montrer l'existence de solutions dissipatives sur des intervalles de temps arbitrairement longs, ce qui n'est, en général, pas possible pour les solutions lisses;
ii) le concept est très utile quand on veut montrer que le SELC peut être dérivé, en un certain sens d'un autre système d'EDP plus compliqué (ou plus "fondamental") en passant à la limite sur de petits paramètres. Il y a beaucoup

d'exemples de cette nature qui ont été établis dans le dernier quart de siècle. (Passages de Navier-Stokes et Boltzmann vers Euler -dans le cas incompressible- respectivement par Pierre-Louis Lions et Laure Saint-Raymond, passage d'Euler vers sa limite "hydrostatique", passage de Schrödinger -non linéaire- vers Euler, etc...etc...).

On peut aussi aller dans l'autre sens, en dérivant des équations plus simples à partir de SELC. On peut, par exemple, passer des équations d'Euler -dans le cas isotherme- vers l'équation de la chaleur, comme on l'avait montré heuristiquement dans le chapitre introductif.

7.4 Retour à l'équation de la chaleur

Dans le chapitre introductif, on a expliqué (formellement) comment l'équation de la chaleur pouvait s'obtenir comme équation "asymptote" de l'équation d'Euler dans le cas compressible isotherme, au prix du changement de variable temporel $t \to t^2$. Plus précisément, on obtient le système "non autonome" (i.e. dépendant explicitement du temps)

$$\partial_t \rho + \nabla \cdot (\rho v) = 0,$$

$$2t\,[\partial_t(\rho v) + \nabla \cdot (\rho v \otimes v)] + \rho v + \nabla \rho = 0,$$

et l'équation de la chaleur (dans le cas de solutions positives ou nulles) est obtenue asymptotiquement en négligeant le terme linéaire en t, ce qui donne

$$\rho v + \nabla \rho = 0$$

et, donc,

$$\partial_t \rho = \Delta \rho.$$

Pour l'équation d'Euler, on a, comme entropie convexe,

$$\mathcal{E}(U) = \frac{|q|^2}{2\rho} + \rho \log \rho - \rho, \quad U = (\rho, \rho v)$$

et donc, si on se place pour simplifier sur le cube périodique $\mathbb{T}^d$, on a pour toute solution lisse

$$\frac{d}{dt} \int_{\mathbb{T}^d} (\rho \log \rho - \rho + \frac{|q|^2}{2\rho}) dx = 0.$$

Après le changement de variable temporel, cette conservation devient

$$\frac{d}{dt} \int_{\mathbb{T}^d} (\rho \log \rho - \rho) dx + 2t \frac{d}{dt} \int_{\mathbb{T}^d} \frac{|q|^2}{2\rho} dx = - \int_{\mathbb{T}^d} \frac{|q|^2}{\rho} dx.$$

Asymptotiquement (en négligeant le terme linéaire en t), on retrouve la relation de "dissipation de l'entropie" bien connue pour l'équation de la chaleur,

$$\frac{d}{dt}\int_{\mathbb{T}^d}(\rho\log\rho-\rho)dx=-\int_{\mathbb{T}^d}\frac{|q|^2}{\rho}dx,$$

ou encore, puisque $q=-\nabla\rho$:

$$\frac{d}{dt}\int_{\mathbb{T}^d}(\rho\log\rho-\rho)dx=-\int_{\mathbb{T}^d}\frac{|\nabla\rho|^2}{\rho}dx,$$

le second membre étant souvent appelée "information de Fisher". Cette relation entre entropie, information de Fisher et équation de la chaleur (qui en définitive remonte -encore- aux équations d'Euler !) a joué un rôle important dans la "théorie du transport optimal" (évoquée dans un chapitre précédent et qui fera l'objet d'un cours FIMFA de M2 par François Bolley), en particulier dans les contributions de C. Villani.
On peut élaborer ces idées pour donner une définition originale de l'équation de la chaleur, développée par Ambrosio, Gigli et Savaré dans les 10 à 15 dernières années. On peut retrouver leur concept sous une forme légèrement différente en raisonnant comme suit. On considère un couple (ρ,q) avec $\rho>0$, $q=\rho v$, solution lisse (pour simplifier la discussion) de l'équation de continuité

$$\partial_t\rho+\nabla\cdot(\rho v)=0,$$

qui est linéaire en $U=(\rho,q)$. On ne suppose rien d'autre pour l'instant. Calculons l'évolution de l'entropie

$$\frac{d}{dt}\int_{\mathbb{T}^d}(\rho\log\rho-\rho)dx=\int_{\mathbb{T}^d}\log\rho\;\partial_t\rho dx=-\int_{\mathbb{T}^d}\log\rho\nabla\cdot(\rho v)dx$$
$$=\int_{\mathbb{T}^d}\nabla\rho\cdot v dx,$$

ce qu'on peut aussi écrire

$$\frac{d}{dt}\int_{\mathbb{T}^d}(\rho\log\rho-\rho)dx+\int_{\mathbb{T}^d}\frac{|q|^2+|\nabla\rho|^2}{2\rho}dx=\int_{\mathbb{T}^d}\frac{|q+\nabla\rho|^2}{2\rho}dx.$$

En intégrant en temps sur $[0,t]$ et notant ρ_0 la valeur initiale de ρ, on déduit

$$\int_{\mathbb{T}^d}(\rho\log\rho-\rho)(t,x)dx+\int_0^t\int_{\mathbb{T}^d}\frac{|q|^2+|\nabla\rho|^2}{2\rho}(t',x)dxdt'=$$
$$\int_{\mathbb{T}^d}(\rho_0\log\rho_0-\rho_0)(x)dx+\int_0^t\int_{\mathbb{T}^d}\frac{|q+\nabla\rho|^2}{2\rho}(t',x)dxdt'.$$

C'est alors qu'on s'aperçoit que $q+\nabla\rho=0$, et donc $(\rho,\rho v)$ est solution de l'équation de la chaleur, si et seulement si on a l'inégalité, pour presque tout $t\geq 0$:

$$\int_{\mathbb{T}^d}(\rho\log\rho-\rho)(t,x)dx+\int_0^t\int_{\mathbb{T}^d}\frac{|q|^2+|\nabla\rho|^2}{2\rho}(t',x)dxdt'\leq\int_{\mathbb{T}^d}(\rho_0\log\rho_0-\rho_0)(x)dx.$$

Ainsi, pour l'équation de la chaleur, écrite sous la forme

$$\partial_t \rho + \nabla \cdot q = 0, \quad q = -\nabla \rho,$$

on obtient le concept suivant de solutions, au moins dans le cas des solutions $\rho \geq 0$:

Definition 43. *On dit que $\rho \in C^0_{w*}([0,T], C^0(\mathbb{T}^d)')$ est solution "dissipative" de donnée initiale ρ_0 si elle vérifie $\rho(0,\cdot) = \rho_0$, et s'il existe $q \in C^0([0,T] \times \mathbb{T}^d; \mathbb{R}^d)'$ telle que*

$$\partial_t \rho + \nabla \cdot q = 0$$

(au sens des distributions) et si elle vérifie pour tout $t \in [0,T]$, l'inégalité

$$\int_{\mathbb{T}^d} (\rho \log \rho - \rho)(t,\cdot) + \int_{[0,t]\times\mathbb{T}^d} \frac{|\nabla \rho|^2 + |q|^2}{2\rho} \leq \int_{\mathbb{T}^d} (\rho_0 \log \rho_0 - \rho_0).$$

Bien entendu, dans cette définition, il convient d'écrire plus précisément l'inégalité qui a priori n'est pas clairement définie en supposant seulement

$$\rho \in C^0_{w*}([0,T], C^0(\mathbb{T}^d)'), \quad q \in C^0([0,T] \times \mathbb{T}^d; \mathbb{R}^d)'.$$

[Notons qu'on travaille ici avec l'espace de mesures $C^0(\mathbb{T}^d)'$ plutôt qu'avec l'espace $L^2(\mathbb{T}^d)$ comme on l'a fait pour les SELC.]
Il s'agit d'abord d'écrire les fonctions convexes $\rho \log \rho - \rho$ et $q^2/(2\rho)$ comme transformées de Legendre-Fenchel. On rappelle notamment que

$$\sup_{u \in \mathbb{R}} u\rho - \exp(u)$$

$$= \rho \log \rho - \rho, \quad \text{si } \rho > 0, \quad 0 \quad \text{si } \rho = 0 \ \text{ et } \ +\infty \ \text{ si } \rho < 0, \text{ alors que}$$

$$\sup\{a\rho + A \cdot q \ ; \quad 2a + |A|^2 \leq 0, \quad a \in \mathbb{R}, A \in \mathbb{R}^d\}$$

$$= |q|^2/(2\rho) \quad \text{si } \rho > 0, \ 0 \text{ , si } \rho = 0 \text{ et } q = 0, \text{ et enfin } \ +\infty \ \text{ sinon.}$$

Ainsi on peut réécrire plus précisément l'inégalité, à l'aide de fonctions test (et d'une intégration par partie pour transformer $\nabla \rho$: exercice !). On exigera donc que, pour tout $t \in [0,T]$, pour toute fonction $u \in C^\infty(\mathbb{T}^d)$ et toute fonction lisse

$$(t,x) \in \mathbb{R}\times\mathbb{T}^d \to (a,b,A,B)(t,x) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d \quad \text{t.q.} \quad 2a+|A|^2 \leq 0, \quad 2b+|B|^2 \leq 0,$$

on a, pour tout $t \in [0,T]$,

$$\int_{\mathbb{T}^d} \Big(u(x)\rho(t,dx) - e^{u(x)}dx \Big) + \int_0^t \int_{\mathbb{T}^d} (a(s,x) + b(s,x) - (\nabla \cdot B)(s,x))\rho(s,dx)ds$$

$$+ \int_{[0,t]\times\mathbb{T}^d} A(s,x) \cdot q(ds,dx) \leq \int_{\mathbb{T}^d} (\rho_0 \log \rho_0 - \rho_0),$$

ce qui est évidemment assez lourd mais est au moins clairement défini quand

$$\rho \in C^0_{w*}([0,T], C^0(\mathbb{T}^d)'), \quad q \in C^0([0,T] \times \mathbb{T}^d; \mathbb{R}^d)'.$$

Comme pour les SELC, l'existence de solutions dissipatives globales en temps $t \geq 0$ est relativement aisée à prouver (au prix, malgré tout, d'une certaine technicité). En revanche, contrairement à se qui se passe dans le cas des SELC, la définition de solution dissipative pour l'équation de la chaleur conduit assez facilement à leur unicité pour une donnée initiale ρ_0 fixée. L'idée est qu'en prenant l'inégalité sous sa forme initiale, on voit que la demi-somme de deux solutions dissipatives distinctes est encore solution dissipative, pour la même donnée initiale bien entendu. Comme l'entropie de Boltzmann $\rho \to \rho \log \rho - \rho$ est *strictement* convexe la demi-somme va vérifier l'inégalité *stricte*. Or, pour une raison qu'il serait trop long d'expliquer, l'inégalité stricte est en fait impossible. D'où contradiction. Observons aussi que, dans la formulation dissipative, la linéarité de l'équation de la chaleur n'est pas du tout évidente. Il est bon d'observer que cette formulation peut se généraliser (à l'aide de techniques de "transport optimal", déjà mentionnées dans un chapitre précédent, au sujet de l'équation de Monge-Ampère) à des espaces métriques très généraux (bien au delà de l'espace euclidien $\mathbb{R}^d$ et même des variétés riemanniennes), comme l'ont fait Luigi Ambrosio, Nicola Gigli et Giuseppe Savaré (cf. Inventiones 2014). Dans certains cas, l'équation de la chaleur (ainsi définie) peut même cesser d'être linéaire... de quoi faire retourner Fourier dans sa tombe !

8 Lois de conservation augmentés avec entropie convexe

Nous couvrirons deux exemples : celui de l'électromagnétisme non-linehaire de Max Born et Leopold Infeld [34] puis celui des surfaces extrémales de dimension quelconque dans l'espace de Minkowski, au moins celles qui s'écrivent comme des graphes, traité en toute généralité par Duan dans [119]. En termes d'applications, ces exemples sont bien connus des spécialistes de physique des hautes énergies (théorie des cordes et des "Dirichlet-branes") [205]. Dans le second cas, on insistera sur le cas plus facile de la co-dimension un où les calculs sont nettement plus explicites et peuvent être facilement exploités, suivant [68], pour traiter les mouvements par courbure moyenne comme on le verra dans un chapitre ultérieur. Dans les deux cas, il s'agit de système de lois de conservation du premier ordre d'origine variationnelle avec des "entropies" (il faudrait plutôt parler d'énergies) non-convexes mais en réalité polyconvexes. Suivant une idée déja introduite par Qiu dans le cadre de l'élasticité non-linéaire [207], on augmente le système en ajoutant des variables supplémentaires pour lesquelles on peut dériver des lois de conservation additionnelles et on obtient une nouvelle "entropie" qui devient convexe par rapport à l'ensemble des nouvelles variables. Dans les deux cas considérer, on parvient même à réécrire le système "augmenté" de lois de conservation sous une forme non divergentielle (i.e. "non conservative" pour reprendre le jargon des lois de conservation) spectaculairement simple :

$$\partial_t U_\alpha + A_\alpha^{i\beta\gamma} U_\gamma \partial_i U_\beta = 0,$$

qu'on a écrite "en coordonnées" (avec sommation implicite sur les indices répétés). Il s'agit donc de généralisations (non triviales) de la fameuse équation de Burgers

$$\partial_t u + u\partial_x u = 0,$$

(sans viscosité). Les solutions $U = U(t,x)$ sont à valeurs dans $\mathbb{R}^m$ et la variable d'espace x appartient à $\mathbb{R}^d$. Les coefficients constants $A_\alpha^{i\beta\gamma}$ ont la propriété que, pour chaque indice $i = 1, \cdots, d$ et $\gamma = 1, \cdots, m$, fixés, la matrice carrée $m \times m$: $(A_\alpha^{i\beta\gamma})$ est symétrique en α, β. Ceci suffit, sans le moindre effort supplémentaire, pour prouver que le problème de Cauchy est bien posé, en temps petit dans tout espace de Sobolev $H^s(\mathbb{R}^d)$ qui s'injecte continument dans C^1, i.e. pour tout $s > 1 + d/2$.

8.1 The Born-Infeld equations

In 1934, Max Born and Leopold Infeld introduced a non-linear correction of the classical Maxwell model. This amounts to finding critical points (with respect to compactly supported perturbations)

$$(t,x) \in \mathbb{R}^{1+3} \to (E,B)(t,x) \in \mathbb{R}^3 \times \mathbb{R}^3,$$

of the following action

$$A_\lambda[E,B] = \int\int (1 - \sqrt{1 + \lambda^{-2}(B^2 - E^2) - \lambda^{-4}(B\cdot E)^2})\ \ dxdt$$

where $\lambda > 0$ is a physical constant (the "absolute field"), under constraints

$$\nabla \cdot B = 0, \quad \partial_t B + \nabla \times E = 0.$$

In the "low-field" limit $\lambda \to \infty$, the classical Maxwell model is recovered

$$\lambda^2 A_\lambda[E,B] \sim \frac{1}{2}\int (E^2 - B^2)\ \ dxdt$$

leading to the famous (homogeneous) Maxwell equations

$$\partial_t B + \nabla \times E = 0, \;\; \partial_t E = \nabla \times B, \;\; \nabla \cdot B = \nabla \cdot E = 0.$$

8.1.1 The electrostatic case

The electrostatic case is consistently obtained by canceling the magnetic field **B**:

$$A_\lambda[E,0] = \int\int (1 - \sqrt{1 - \lambda^{-2}E^2})\ \ dxdt$$

under constraint

$$\nabla \times E = 0.$$

So, the constant $\lambda > 0$ just appears as the maximal possible electrostatic field in the theory (just like 1 is the maximal possible velocity in Special Relativity). This was Max Born's original idea.

8.1.2 Remark: a more general and geometric definition

For a general $\mathbf{1+d}$ dimensional Lorentzian manifold with metric $g_{ij}dx^i dx^j$ the BI model involves a closed 2-form $\mathcal{B}_{ij}dx^i \wedge dx^j$ and the Born-Infeld Action now reads

$$A_\lambda[g,\mathcal{B}] = \int (\sqrt{-\det g} - \sqrt{-\det(g+\lambda\mathcal{B})}).$$

Notice that this Action is "fully covariant", i.e. invariant as g and $\mathcal{B}$ are deformed by any space-time diffeomorphism. (Indeed, there is an exact compensation between the determinant and the modifications brought to $g_{ij}dx^i dx^j$ and $B_{ij}dx^i \triangle up dx^j$ by any diffeonorphism

$$x = (x^0,\cdots,x^d) \in R^{1+d} \to \Phi(x) \in R^{1+d}.$$

Of course, in the special case $d = 3$, $g = \text{diag}(-1,1,1,1)$,one may recover (through an elementary but interesting calculation, involving elementary linear algebra and properties of $4\times$ skew symmetric matrices) the previous formulae introduced in 1934 in the special case of the standard 1+3 Minkowski space.

8.1.3 Remark: high-field limit of the Born-Infeld model and Magnetohydrodynamics

The original Born-Infeld model

$$A_\lambda[E,B] = \int\int (1 - \sqrt{1+\lambda^{-2}(B^2-E^2) - \lambda^{-4}(B\cdot E)^2})\ \ dxdt$$

$$\nabla\cdot B = 0, \quad \partial_t B + \nabla\times E = 0$$

admits an interesting "high-field" limit obtained as $\lambda \to \mathbf{0}$, namely, at least formally,

$$\lambda A_\lambda[E,B] \sim -\int\int \sqrt{B^2-E^2}\ \ dxdt$$

under the additional pointwise constraint $\mathbf{E}\cdot\mathbf{B} = \mathbf{0}$. This pointwise constraint $\mathbf{E}\cdot\mathbf{B} = \mathbf{0}$ is equivalent to $\mathbf{E} = \mathbf{B}\times\mathbf{v}$ for some new field $\mathbf{v} = \mathbf{v}(\mathbf{t},\mathbf{x})$. This leads to

$$\lambda A_\lambda[E,B] \sim -\int\int \sqrt{B^2(1-v^2)+(B\cdot v)^2}\ \ dxdt$$

with differential constraints

$$\nabla\cdot B = 0, \quad \partial_t B + \nabla\times(B\times v) = 0$$

which can be interpreted as the "induction equation" in ideal Magnetohydrodynamics, where B and v may be seen respectively as the magnetic field and the velocity field of a charged fluid.

8.1.4 The Born-Infeld equations in Hamiltonian form

After normalization $\lambda = 1$, written in Hamiltonian form, the Born-Infeld equations read

$$\partial_t B + \nabla \times \left(\frac{B \times (D \times B) + D}{\sqrt{1 + D^2 + B^2 + (D \times B)^2}}\right) = 0, \quad \nabla \cdot B = 0,$$

$$\partial_t D + \nabla \times \left(\frac{D \times (D \times B) - B}{\sqrt{1 + D^2 + B^2 + (D \times B)^2}}\right) = 0, \quad \nabla \cdot D = 0.$$

As shown by Speck [222], using Klainerman's null forms, global smooth solutions to the initial value problem have been proven to uniquely exist for small localized initial conditions. We are going to follow a very different way to analyse the Born-Infeld equations, by augmenting the system and finding a suitable convex "entropy function".

8.1.5 The energy-momentum conservation laws

By Noether's theorem, since the Born-Infeld Action is manifestly invariant under time and space translations in the Minkowski space $\mathbb{R}^{1+3}$, we expect four extra conservation laws. There calculation is elementary but not completely obvious:

$$\partial_t Q + \nabla \cdot \left(\frac{Q \otimes Q - B \otimes B - D \otimes D}{h}\right) = \nabla\left(\frac{1}{h}\right), \qquad \partial_t h + \nabla \cdot Q = 0$$

for the energy and momentum fields

$$h = \sqrt{1 + D^2 + B^2 + (D \times B)^2}, \quad Q = D \times B.$$

8.1.6 The augmented Born-Infeld system

Following [55] we define the 10 by 10 *augmented* Born-Infeld system (ABI) as the original BI system **augmented** by the 4 energy-momentum conservation laws

$$\partial_t B + \nabla \times \left(\frac{B \times Q + D}{h}\right) = \partial_t D + \nabla \times \left(\frac{D \times Q - B}{h}\right) = 0$$

$$\partial_t Q + \nabla \cdot \left(\frac{Q \otimes Q - B \otimes B - D \otimes D}{h}\right) = \nabla\left(\frac{1}{h}\right), \qquad \partial_t h + \nabla \cdot Q = 0$$

while disregarding the original algebraic constraints

$$h = \sqrt{1 + D^2 + B^2 + (D \times B)^2}, \quad Q = D \times B,$$

which define a 6 dimensional algebraic submanifold in the space $(h, Q, D, B) \in R^{10}$ that we call the "BI manifold".

8.1.7 The ABI system in non-conservative variables

Here, our analysis follows [79] rather than [55]. Indeed, the augmented BI system looks even simpler in so-called "non-conservative variables"

$$b = B/h, \quad d = D/h, \quad v = Q/h, \quad \tau = 1/h$$

Namely

$$\partial_t b + (v \cdot \nabla)b - (b \cdot \nabla)v + \tau\nabla \times d = 0$$
$$\partial_t d + (v \cdot \nabla)d - (d \cdot \nabla)v - \tau\nabla \times b = 0$$
$$\partial_t v + (v \cdot \nabla)v - (b \cdot \nabla)b - (d \cdot \nabla)d - \tau\nabla\tau = 0$$
$$\partial_t \tau + (v \cdot \nabla)\tau - \tau\nabla \cdot v = 0$$

This turns out to be just a symmetric system with purely quadratic nonlinearities! In some, a generalization of the inviscid Burgers equation, of form

$$\partial_t U_\alpha + A_\alpha^{i\beta\gamma} U_\gamma \partial_i U_\beta = 0,$$

written "in coordinates" (with implicit summation on repeated indices), where $U = U(t,x) \in \mathbb{R}^{10}$ and, for each fixed indices $i = 1, \cdots, 3$ and $\gamma = 1, \cdots, 10$, the 10×10 matrices $(A_\alpha^{i\beta\gamma})$ are symmetric in α, β. Also observe that there is no limitation of range for the variables $U = (n, d, v, \tau)$ in the space $\mathbb{R}^{10}$. (In particular it makes sense to consider negative or null values of τ, which is not possible in the conservative formulation of the ABI system since $\rho = 1/\tau$. This is a remarkable advantage of the non-conservative version! Of course, we don't make any comment on the possible physical meaning of considering negative values of τ!) Concerning the BI manifold, its expression in terms of non-conservative variables is even simpler. We get the following algebraic (quadratic) 6-dimensional submanifold of $\mathbb{R}^{10}$:

$$\text{NCBIM} \;\; \tau^2 + b^2 + d^2 + v^2 = 1, \;\; \tau v = d \times b.$$

(Notice that we may consider both positive and negative values of τ in this definition!)

So, we obtain, essentially for free, the following result

Theorem 44. *The non-conservative augmented Born-Infeld (NCABI) system is locally well-posed in any Sobolev space $H^s(\mathbb{R}^3)$ continuously imbedded in C^1 (namely, for any $s > 5/2$). In addition the non-conservative Born-Infold manifold is preserved under evolution.*

Because of the preservation of the manifold, we have immediately, without any further analysis, obtained the local well-posedness of the orginal Born-Infeld equations. Of course, the analysis provided by Speck [222] is much more sophisticated and leads to a *global* existence and uniqueness result of smooth solutions to the expanse of assuming to initial conditions to be small and localized, which is in no way needed in our cruder analysis. An interesting open question is the possible global existence of smooth solutions not only for the original BI system but also for its augmented version.

8.1.8 Remark: reduced versions of the NCABI system: motion of strings and photons

It is perfectly consistent to assume $\tau = 0$, $d = 0$ in the non-conservative augmented BI (NCABI) system. We then get a reduced system which describes a continuum of vibrating strings

$$\partial_t b + (v \cdot \nabla) b - (b \cdot \nabla) v = 0, \quad \partial_t v + (v \cdot \nabla) v - (b \cdot \nabla) b = 0$$

The corresponding BI manifold $b^2 + v^2 = 1$, $v \cdot b = 0$ corresponds to relativistic strings, like in "classical" String Theory (i.e. without quantization). We may further consistently assume $b = 0$ in the NCABI and get $\partial_t v + (v \cdot \nabla) v = 0$ with reduced BI-manifold $v^2 = 1$ which describes the motion of (classical) massless particles moving at the speed of light (e.g. photons).

8.1.9 First appearance of convexity in the augmented Born-Infeld system

Let us now go back to the 10×10 augmented ABI system in conservative form. Surprisingly enough, the augmented system, as shown in [55], admits an extra conservation law, namely

$$\partial_t \eta + \nabla \cdot \omega = 0, \quad \eta = \frac{1 + D^2 + B^2 + Q^2}{h}, \quad \omega = \omega(h, Q, D, B)$$

where η is a strictly *convex* function and the "entropy flux" can be explicitly computed. This makes the ABI system an example of entropic system of conservation laws (ESCL), for which we can use all the concepts of "relative entropy method" and "dissipative solutions".

8.1.10 Remark: Galilean invariance of the augmented Born-Infeld system

The ABI system looks pretty much like classical MHD equations and enjoys an astonishing *classical* Galilean invariance, under the transform

$$(t, x) \to (t, x + W\ t), \quad (h, Q, D, B) \to (h, Q - hU, D, B)$$

for any constant speed $W \in R^3$! This looks contradictory with the definite Lorentzian origin of the Born-Infeld system. However, there is no contradiction since those Galilean transforms are *incompatible* with the Born-Infeld manifold, where Q is algebraically slaved by B and D through $Q = D \times B$! Moreover, we conjecture that this amazing property characterizes the Born-Infeld model among all alternative Electromagnetic theories, including ...Maxwell's one!

8.1.11 Second appearance of convexity in the augmented Born-Infeld system

The 10×10 ABI (augmented Born-Infeld) system is *linearly degenerate* (in the sense of Lax [106]) and enjoy an interesting stability under weak-* convergence. More precisely:

Theorem 45. *Each weak-* limit of uniformly bounded sequences in L^∞ of smooth solutions depending on one space variable of the ABI system are still solutions of the ABI system.*

This follows from a straightforward application of the Murat-Tartar 'div-curl' lemma [196, 225]. This suggests that the convex hull of the BI manifold might be a natural completed configuration space for the Born-Infeld theory. However, this is not so clear, as pointed out to the author by Felix Otto, since one has to take into account the differential constraints $\nabla \cdot D = \nabla \cdot B = 0$. Anyway, as shown in [55], the convex hull has full dimension in $\mathbb{R}^{10}$ and has been explicitly computed by Serre [215] and is defined by the single inequality

$$h \geq \sqrt{1 + D^2 + B^2 + Q^2 + 2\sqrt{|P - D \times B|^2 + (B \cdot P)^2 + (D \cdot P)^2}}$$

Moreover Müller and Palombaro [195], using convex integration theory, have proven that the differential constraints $\nabla \cdot D = \nabla \cdot B = 0$ are not an obstruction to the conjecture.
On the convexified BI manifold, defined by Serre's inequality, we have the following properties:
1) The electromagnetic field (D, B) and the 'density and momentum' fields (h, Q) can be chosen *independently* of each other at initial time, provided they satisfy Serre's inequality
2) The augmented BI system can be interpreted (in MHD style) as the coupling of an electromagnetic field with a fluid

$$\partial_t B + \nabla \times (\frac{B \times Q + D}{h}) = \partial_t D + \nabla \times (\frac{D \times Q - B}{h}) = 0$$

$$\partial_t Q + \nabla \cdot (\frac{Q \otimes Q - B \otimes B - D \otimes D}{h}) = \nabla(\frac{1}{h}), \quad \partial_t h + \nabla \cdot Q = 0.$$

(while the original Born-Infeld model is purely electromagnetic, without any interaction with matter).
3) 'Matter' may exist without electromagnetic field, in the case when $B = D = 0$, which leads to the so-called "Chaplygin gas" (which has been advocated as a possible model for "dark energy" or "vacuum energy") with an usual speed of sound c, namely $c = 1/h$,

$$\partial_t Q + \nabla \cdot (\frac{Q \otimes Q}{h}) = \nabla(\frac{1}{h}), \quad \partial_t h + \nabla \cdot Q = 0$$

4) 'Moderate' Galilean transforms are allowed

$$(t, x) \to (t, x + U\, t), \quad (h, Q, D, B) \to (h, Q - hU, D, B)$$

(which is impossible on the original BI manifold). As a matter of fact, this seems to be a general feature of Special Relativity under weak completion (cf. "subrelativistic" conditions, as discussed in [56, 21].

9 Convex entropic formulation of some degenerate parabolic systems

As we have already seen in the previous chapter, entropy methods can be used in a fruitful way already at the level of the simple linear heat equation. In the present chapter, we would like to address more sophisticated examples, typically mean curvature flows of various co-dimensions. Our main idea here is to address those parabolic systems that we can derive from systems of conservation laws with convex entropy and to transfer the rather comprehensive theory we had obtained for the later, as explained in the previous chapter. This is precisely the case of mean curvature flows as we will see. Our main trick to derive parabolic systems from systems of first order conservation laws is extremely simple, although not yet used, at least for evolution PDEs, in the literature, to the best of our knowledge. It amounts to performing a quadratic change of time near $t = 0$ and, then, neglecting the higher order terms. Let us emphasise that this is just an algebraic trick and there is no analysis involved at this level.

9.0.1 Derivation of the heat and the porous media equations from the Euler equations of compressible fluids

As a leitmotiv, we go back to the Euler equations, this time for *compressible* fluids, as written by Euler in 1755-57(i.e. without thermodynamics nor energy equation: they are frequently called "isentropic Euler equations"):

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \partial_t(\rho v) + \nabla \cdot (\rho v \otimes v) = -\nabla p$$

where $(\rho, p, v) \in \mathbb{R}^{1+1+3}$ are the density, pressure and velocity fields of a fluid and p is assumed to be a given function of ρ. Let us perform the quadratic change of time (QCT)

$$t \to \theta = t^2/2, \quad (\rho, v)(t, x) \to (\rho(\theta, x), \theta' v(\theta, x)), \quad \theta' = \frac{d\theta}{dt} = t$$

Let us perform the following quadratic change of time (QCT)

$$\tilde{\rho}(t, x) = \rho(\theta, x), \quad \tilde{v}(t, x) = \theta' v(\theta, x), \quad \theta = \theta(t) = t^2/2 \quad \theta' = \frac{d\theta}{dt} = t$$

(so that $\tilde{v}(t, x)dt = v(\theta, x)d\theta$). We get:

$$\partial_t \tilde{\rho} + \nabla \cdot (\tilde{\rho}\tilde{v}) = 0 \quad \to \quad \theta' \partial_\theta \rho + \theta' \nabla \cdot (\rho v) = 0$$

$$\partial_t(\tilde{\rho}\tilde{v}) + \nabla \cdot (\tilde{\rho}\tilde{v} \otimes \tilde{v}) = -\nabla p(\tilde{\rho}) \quad \to$$

$$\theta" \rho v + (\theta')^2 \partial_\theta(\rho v) + (\theta')^2 \nabla \cdot (\rho v \otimes v) = -\nabla p(\rho)$$
$$\rightarrow \quad \rho v + 2\theta \partial_\theta(\rho v) + 2\theta \nabla \cdot (\rho v \otimes v) = -\nabla p(\rho)$$

So, after the quadratic change of time, the Euler equations become

$$\partial_\theta \rho + \nabla \cdot (\rho v) = 0, \ \rho v + 2\theta[\partial_\theta(\rho v) + \nabla \cdot (\rho v \otimes v)] = -\nabla p(\rho)$$

Notice that the continuity equation has stayed unchanged. (Actually, this was the main purpose of the different rescaling of variables ρ and v.) The new system of evolution PDEs is no longer "autonomous": it depends *explicitly* on the new time variable θ, actually in a very simple, linear, way. So we may consider two asymptotic regimes, according to the size of θ. For very large θ, we just obtain the so-called "pressureless Euler" equations:

$$\partial_\theta \rho + \nabla \cdot (\rho v) = 0, \ \rho v + \partial_\theta(\rho v) + \nabla \cdot (\rho v \otimes v) = 0,$$

which is just a degenerate (but tricky!) version of the Euler equations. We are much more interested in the second regime when θ is very small. Then, we obtain the so-called "porous media equation" (PME)

$$\partial_\theta \rho + \nabla \cdot (\rho v) = 0, \quad \rho v = -\nabla p,$$

or, in short,

$$\text{PME}: \qquad \partial_\theta \rho = \Delta(p(\rho)),$$

including the heat equation in the special ("isothermal") case $p(\rho) = \rho$. So, the QCT has clearly introduced a change of type in the equations, since we have moved from the hyperbolic, first order, setting of the Euler equations to the parabolic, second order in space, setting of the heat and the porous medium equations.

9.0.2 From dynamical systems to gradient flows

Actually, the quadratic change of time method (QCTM) can already be applied to the simple dynamical system

$$\frac{d^2X}{dt^2} = -(\nabla\varphi)(X),$$

by setting

$$X(t) = Y(\theta), \quad \theta = t^2/2 \quad \theta' = \frac{d\theta}{dt} = t.$$

This leads to

$$\frac{dX}{dt} = \theta' \frac{dY}{d\theta}, \quad -(\nabla\varphi)(Y(\theta)) = \frac{d}{dt}(\theta' \frac{dY}{d\theta}) = \theta" \frac{dY}{d\theta} + (\theta')^2 \frac{d^2Y}{d\theta^2}$$

and thus

$$\frac{dY}{d\theta} + 2\theta \frac{d^2Y}{d\theta^2} = -(\nabla\varphi)(Y).$$

For large θ, we get the purely inertial motion governed by:

$$\frac{d^2Y}{d\theta^2} = -(\nabla\varphi)(Y),$$

while, for small θ, we rather get the so-called "gradient flow" regime with:

$$\frac{dY}{d\theta} = -(\nabla\varphi)(Y).$$

Remark:
The quadratic rescaling $\theta = t^2/2$ perfectly fits with Galileo's experiment: a rigid ball descends a rigid ramp of constant slope, with zero initial velocity and constant acceleration G, reaching position $X(t) = x_0 + Gt^2/2 = x_0 + G\theta = Y(\theta)$ at time t. So, Y is just a linear function of the rescaled time θ!

$$\frac{dY}{d\theta} + 2\theta\frac{d^2Y}{d\theta^2} = G$$

but also *simultaneously*

$$\frac{dY}{d\theta} = G, \quad \frac{d^2Y}{d\theta^2} = 0,$$

i.e. *both* the gradient flow and the inertial regimes.
End of remark.

set a collection of little bells along the ramp according to a parabolic spacing $(1, 4, 9, 16, 25...)$. Then, the ball will ring the bells according to a linear progression. $T, 2T, 3T, 4T, 5T, ...$, which can be checked with the help of a (sufficiently accurate clock).

For the original dynamical system,

$$\frac{d^2X}{dt^2} = -\nabla\varphi(X),$$

we get the usual conservation of energy

$$\frac{d}{dt}[\frac{1}{2}|\frac{dX}{dt}|^2 + \varphi(X)] = 0$$

For the time-rescaled version $Y(\theta) = X(t)$, $\theta = t^2/2$, we find

$$\frac{d}{d\theta}[\varphi(Y)] + \theta\frac{d}{d\theta}|\frac{dY}{d\theta}|^2 = -|\frac{dY}{d\theta}|^2$$

In the asymptotic regime when θ is very small, we recover the gradient flow

$$\frac{dY}{d\theta} = -\nabla\varphi(Y)$$

and the classical "$energy - dissipation$" relation

$$\frac{d}{d\theta}[\varphi(Y)] + = -|\frac{dY}{d\theta}|^2.$$

We may compare, for short times, X solution of the original equation, with zero initial velocity, to Y solution of the gradient flow

$$\frac{d^2X}{dt^2} = -\nabla\varphi(X), \;\; X'(0) = 0, \;\; \frac{dY}{d\theta} = -\nabla\varphi(Y), \;\; Y(0) = X(0).$$

Under strong convexity and smoothness assumptions on φ, Assuming $\mathbf{spect}(\mathbf{Hess})(\varphi) \subset [\mathrm{r}, 1/\mathrm{r}]$ for some constant $r > 0$, we may easily prove

$$|X(t) - Y(t^2/2)|^2 + |\frac{dX}{dt}(t) - t\frac{dY}{d\theta}(t^2/2)|^2 \leq t^4\exp(t^2c)c.$$

by monitoring the "relative energy"

$$\frac{1}{2}|\frac{dX}{dt} - t\frac{dY}{d\theta}|^2 + \varphi(X) - \varphi(Y) - \nabla\varphi(Y)\cdot(X-Y),$$

which is just obtained (as a "relative entropy") by substracting from the energy of X what we obtain by expanding linearly the energy in X about Y. Notice that constant c depends only on r and on Y.

9.0.3 Quadratic change of time for geometric equations

We are now going to get some optimal transport gradient flows (typically mean curvature flows) from hyperbolic equations (typically geometric wave equations) through the quadratic change of time method. This has been developed for the curve-shortening flow (which is the mean-curvature flow in dimension 1, i.e. in co-dimension $d-1$), with Xianglong Duan [71]. Here, we emphasise the substantially simpler case of mean curvature flow for graphs, with co-dimension one, studied in [68].

Theorem 46. *Through the quadratic change of time method, the nonlinear wave equation, which describes graphs of extremal area in the Minkowski space* $\mathbb{R}^{1+d}$*,*

$$\partial_t(\frac{\partial_t\phi}{R}) = \nabla\cdot(\frac{\nabla\phi}{R}), \quad R = \sqrt{1-\partial_t\phi^2+|\nabla\phi|^2}$$

generates two twin evolution PDEs. The first one is the "arctangential" heat equation $\partial_t D = \Delta(\arctan D)$*, while the second one is just the well known mean curvature flow for graphs*

$$\partial_t\phi = \sqrt{1+|\nabla\phi|^2}\ \ \nabla\cdot\left(\frac{\nabla\phi}{\sqrt{1+|\nabla\phi|^2}}\right).$$

Remark: a "naive" interpretation of the arctangential heat equation in optimal transport terms:

The arctangential flow $\partial_t D = \lambda\Delta(\arctan(D\lambda^{-1})$ (where we have input the scaling parameter $\lambda > 0$) can be easily written in optimal transport style (à la Otto) [201, 202]

$$\partial_t D = \nabla\cdot\left(D\ \ \nabla(\mathcal{F}'(D))\right),$$

where

$$\mathcal{F}(D) = D\log\left(\frac{D}{\sqrt{1+D^2\lambda^{-2}}}\right) - \lambda\arctan(D\lambda^{-1})$$

is the Legendre transform of

$$u \to \lambda\arcsin(\lambda^{-1}e^u)$$

(extended by $+\infty$ for $u > \log\lambda$), which can be seen as a "catastrophic" version of the usual exponential. (N.B. The inverse of this "catastrophic" exponential $u \to \lambda\arcsin(\lambda^{-1}\exp(u))$ can be symmetrized and periodized as $v \to \frac{1}{2}\log(\lambda^2\sin^2(v\lambda^{-1}))$, which, surprisingly enough, also plays a crucial role in the recent theory of "unbalanced optimal transport" [172, 97] ic" exponential function, drawn for different values of parameter λ, and its inverse (after symmetrization and periodization): $v \to \frac{1}{2}\log(\lambda^2 \sin^2(v\lambda^{-1}))$

Gradient flow structure of the arctangential flow

In a less naive and more geometric way, the arctangential heat equation

$$\partial_t D = \Delta(\arctan(D))$$

can be seen à la Otto as the gradient flow of functional

$$D \to \int_{\mathbb{T}^d} \sqrt{1 + D^2}$$

with metric

$$||v||_D^2 = \int_{\mathbb{T}^d} (1 + D^2)^{-1/2} |vD|^2,$$

where tangent vectors $\dot{D}$ at point D are written

$$\dot{D} = -\nabla \cdot (vD).$$

This (hidden) structure is naturally inherited from the nonlinear wave equation

$$\partial_t(\frac{\partial_t \phi}{R}) = \nabla \cdot (\frac{\nabla \phi}{R}), \quad R = \sqrt{1 - \partial_t \phi^2 + |\nabla \phi|^2},$$

through the quadratic change of time method.

Gradient flow structure of the mean curvature flow

In a parallel way, the natural gradient flow structure of the mean curvature equation for graphs is inherited from the same nonlinear wave equation. It corresponds to In a parallel way, the mean curvature flow for graphs can be interpreted as the gradient flow of functional

$$B \to \int_{\mathbb{T}^d} \sqrt{1 + |B|^2}$$

with metric

$$||v||_B^2 = \int_{\mathbb{T}^d} (1 + |B|^2)^{-1/2} (B \cdot v)^2$$

where tangent vectors at point B are written as

$$\dot{B} = -\nabla(B \cdot v).$$

Proof of Theorem 46

We want to derive from the nonlinear wave equation (studied by Lindblad in [173])

$$\partial_t(\frac{\partial_t \phi}{R}) = \nabla \cdot (\frac{\nabla \phi}{R}), \quad R = \sqrt{1 - \partial_t \phi^2 + |\nabla \phi|^2},$$

at once, both the arctangential heat flow

$$\partial_t D = \Delta(\arctan D)$$

the mean curvature flow for graphs

$$\partial_t \phi = \sqrt{1+|\nabla\phi|^2}\ \ \nabla\cdot\left(\frac{\nabla\phi}{\sqrt{1+|\nabla\phi|^2}}\right).$$

Proof/First step.
Here we proceed as we did for the Born-Infeld equations, by introducing a suitable augmented system revealing the hidden convexity structure of the wave equation. More precisely:

Theorem 47. *As $\phi(t,x)$ solves the equation of extremal surfaces in Minkowski's space, then*

$$(D,B,P)=\frac{1}{\sqrt{1-\partial_t\phi^2+|\nabla\phi|^2}}(\partial_t\phi,\nabla\phi,-\partial_t\phi\ \nabla\phi)$$

solves the "entropic" system of conservation laws:

$$\partial_t B+\nabla\left(\frac{P\cdot B-D}{h}\right)=0,\quad \partial_t D+\nabla\cdot\left(\frac{PD-B}{h}\right)=0,$$

$$\partial_t P+\nabla\cdot\left(\frac{P\otimes P+B\otimes B}{h}\right)=\nabla\left(\frac{1+B^2}{h}\right),$$

with

$$h=h(D,B,P)=\sqrt{1+D^2+B^2+P^2}$$

as convex "entropy", which is a strictly convex function of (D,B,P) and obeys an extra conservation law.

Let us postpone the proof of this result for a moment and continue the proof of Theorem 46.

Proof of Theorem 46. */Second step.*
We apply the quadratic change of time method $t\to\theta=t^2/2$ in two different ways. A first possible rescaling is

$$\mathcal{B}(\theta,x)=B(\sqrt{2\theta},x),$$

$$\mathcal{D}(\theta,x)=\frac{D(\sqrt{2\theta},x)}{\sqrt{2\theta}},\quad \mathcal{P}(\theta,x)=\frac{P(\sqrt{2\theta},x)}{\sqrt{2\theta}},$$

requiring initial condition $D=P=0$ at $t=0$, which corresponds to $\partial_t\phi(0,x)=0$ in terms of the solution ϕ to the nonlinear wave equation.
In a somewhat dual way, a second natural change is

$$\mathcal{D}(\theta,x)=D(\sqrt{2\theta},x),$$

$$\mathcal{B}(\theta,x)=\frac{B(\sqrt{2\theta},x)}{\sqrt{2\theta}},\quad \mathcal{P}(\theta,x)=\frac{P(\sqrt{2\theta},x)}{\sqrt{2\theta}},$$

requiring initial condition $B = P = 0$ at $t = 0$, which corresponds to $\nabla\phi = 0$ at $t = 0$ in terms of ϕ.
After performing the change of time $t \to \theta = t^2/2$, we get, in the 1st case, the non automous system:

$$\partial_\theta \mathcal{B} = \nabla\left(\mathcal{D} - \frac{\mathcal{P}\cdot\mathcal{B}}{\mathcal{H}}\right),\ \mathcal{H} = \sqrt{1 + \mathcal{B}^2 + 2\theta(\mathcal{D}^2 + \mathcal{P}^2)},$$

$$\mathcal{D} - \nabla\cdot\left(\frac{\mathcal{B}}{\mathcal{H}}\right) = -2\theta\left(\partial_\theta\mathcal{D} + \nabla\cdot\left(\frac{\mathcal{P}\mathcal{D}}{\mathcal{H}}\right)\right),$$

$$\mathcal{P} + \nabla\cdot\left(\frac{\mathcal{B}\otimes\mathcal{B}}{\mathcal{H}}\right) - \nabla\left(\frac{1+\mathcal{B}^2}{\mathcal{H}}\right)$$

$$= -2\theta\left(\partial_\theta\mathcal{P} + \nabla\cdot\left(\frac{\mathcal{P}\otimes\mathcal{P}}{\mathcal{H}}\right)\right),$$

Neglecting the red terms leads to the mean curvature flow (for graphs), written as an augmented system, in form:

$$\partial_\theta \mathcal{B} = \nabla\left(\mathcal{D} - \frac{\mathcal{P}\cdot\mathcal{B}}{\mathcal{H}}\right),\ \ \mathcal{H} = \sqrt{1 + \mathcal{B}^2}$$

$$\mathcal{D} = \nabla\cdot\left(\frac{\mathcal{B}}{\mathcal{H}}\right),\ \ \mathcal{P} + \nabla\cdot\left(\frac{\mathcal{B}\otimes\mathcal{B}}{\mathcal{H}}\right) = \nabla\left(\frac{1+\mathcal{B}^2}{\mathcal{H}}\right).$$

Symmetrically, the second rescaling leads to the arctangential heat equation and, then, the twin gradient flow structures easily follow.
End of proof.

9.0.4 Proof of Theorem 47

First step : Hamiltonian form of the minimal surface equations.
The non linear wave equation

$$\partial_t\left(\frac{\partial_t\phi}{R}\right) = \nabla\cdot\left(\frac{\nabla\phi}{R}\right), \quad R = \sqrt{1 - \partial_t\phi^2 + |\nabla\phi|^2},$$

is easily obtained by finding critical points ϕ of the Minkowski area of the graph $(t, x) \to (t, x, \phi(t, x))$, namely

$$-\int\int \sqrt{1 - \partial_t\phi^2 + \partial^k\phi\,\partial_k\phi}\;dtdx,$$

under space-time compactly supported perturbations. For the sequel, it is crucial to use the Hamiltonian form of the nonlinear wave equation. For that purpose, we introduce the fields

$$E(t, x) = \partial_t\phi(t, x), \quad B_i(t, x) = \partial_i\phi(t, x),$$

which are linked by the differential constraint $\partial_t B_i = \partial_i E$. Introducing the Lagrangian function

$$L(E,B) = -\sqrt{1 - E^2 + B_k B^k},$$

we look at critical points (E,B) of

$$\int\int L(E(t,x), B(t,x))dtdx$$

under space-time compactly supported perturbations, subject to the differential constraints. In other words, we look for saddle-points (E,B,ψ) of

$$\int\int \left(L(E(t,x), B(t,x)) + \partial_t\psi^i B_i(t,x) - \partial_i\psi^i E(t,x)\right) dtdx$$

where $\psi = \psi(t,x) \in \mathbb{R}^d$ is a Lagrange multiplier for the differential constraint. Independently of the specific definition of L, we may introduce the Hamiltonian H as the partial Legendre-Fenchel transform of the Lagrangian $L(E,B)$ with respect to E,

$$H(D,B) = \sup_{E\in\mathbb{R}} DE - L(E,B)$$

and the corresponding "dual" field

$$D(t,x) = (\frac{\partial L}{\partial E})(E(t,x), B(t,x)).$$

Then, we get, by standard differential calculus, the Hamiltonian formulation

$$\partial_t B_i = \partial_i\left(\frac{\partial H}{\partial D}(D,B)\right), \quad \partial_t D = \partial_i\left(\frac{\partial H}{\partial B_i}(D,B)\right),$$

and, as a consequence, an extra conservation law involving H

$$\partial_t(H(D,B)) + \partial_i(P^i(D,B)) = 0, \quad P^i(D,B) = \left(\frac{\partial H}{\partial D}\frac{\partial H}{\partial B_i}\right)(D,B).$$

In the case of the nonlinear wave equation we get, explicity,

$$H(D,B) = \sqrt{(1 + B_k\ B^k)(1 + D^2)}$$

and, after elementary calculations, deduce

Proposition 48. *The nonlinear wave equation*

$$\partial_t(\frac{\partial_t\phi}{R}) = \nabla\cdot(\frac{\nabla\phi}{R}), \quad R = \sqrt{1 - \partial_t\phi^2 + |\nabla\phi|^2},$$

can be written in Hamiltonian form

$$\partial_t B_i = \partial_i\left(\sqrt{\frac{1 + B_k B^k}{1 + D^2}}D\right), \quad \partial_t D = \partial_i\left(\sqrt{\frac{1 + D^2}{1 + B_k B^k}}\ B^i\right), \tag{34}$$

with the extra-conservation law

$$\partial_t H + \partial_i P^i = 0, \quad H = \sqrt{(1 + B_k\ B^k)(1 + D^2)}, \quad P^i = -DB^i.$$

In addition, (D, B) are related to ϕ by

$$B_i = \partial_i \phi, \quad D = \frac{\partial_t \phi}{\sqrt{1 - \partial_t \phi^2 + \partial^k \phi\ \partial_k \phi}}.$$

Second|; step Construction of an augmented system with convex entropy. Since the Hamiltonian

$$H(D, B) = \sqrt{(1 + B_k\ B^k)(1 + D^2)}$$

is, unfortunately, not a convex function of (D, B), and, therefore the hamiltonian form of the nonlinear wave equation (34) does not belong to our favorite class of systems of entropic system of conservation laws with a convex entropy. However, there is also an extra conservation law for $P = -DB$, namely

$$\partial_t P + \nabla \cdot \left(\frac{P \otimes P + B \otimes B}{h} \right) = \nabla \left(\frac{1 + B^2}{h} \right),$$

where $h = h(D, B, P) = \sqrt{1 + D^2 + B_k B^k + P_k P^k}$ is nothing but $H(D, B)$, written as a function of (D, B, P). We can *add* this new conservation laws to the one we have previously obtained for (D, B), namely

$$\partial_t B + \nabla \left(\frac{P \cdot B - D}{h} \right) = 0, \quad \partial_t D + \nabla \cdot \left(\frac{PD - B}{h} \right) = 0$$

(where we input the new variable h). This allows us, ignoring the algebraic constraint $P = -DB$, to consider (D, B, P), as a solution of an *augmented* system of conservation laws which turns out to enjoy an extra conservation law for the strictly convex "entropy" $h(D, B, P) = \sqrt{1 + D^2 + B_k B^k + P_k P^k}$. Let us just reproduce the detailed calculations provided in the appendix of [68]. Since $P_i = DB_i$, we get

$$\partial_t P_i = -D \partial_t B_i - B_i \partial_t D = T = T4 + T3 + T1 + T2,$$

$$T4 = D \partial_i \left(\frac{B_j P^j}{h} \right), \quad T3 = -D \partial_i \left(\frac{D}{h} \right),$$

$$T1 = B_i \partial_j \left(\frac{D P^j}{h} \right), \quad T2 = -B_i \partial_j \left(\frac{B^j}{h} \right)$$

We have

$$T4 = T4a + T4b, \quad T4a = D B_j \partial_i \left(\frac{P^j}{h} \right), \quad T4b = \frac{D P^j}{h} \partial_i B_j,$$

$$T3 = T3a + T3b, \;\; T3a = -\partial_i\left(\frac{D^2}{h}\right), \;\; T3b = \frac{D}{h}\partial_i D,$$

$$T1 = T1a + T1b, \;\; T1a = \partial_j\left(\frac{B_i D P^j}{h}\right), \;\; T1b = -\partial_j B_i \frac{DP^j}{h},$$

$$T2 = T2a + T2b, \;\; T2a = -\partial_j\left(\frac{B_i B^j}{h}\right), \;\; T2b = \partial_j B_i \frac{B^j}{h}.$$

Since $P_j = -DB_j$, we have

$$T1a = -\partial_j\left(\frac{P_i P^j}{h}\right),$$

$$T4a = -P_j\partial_i\left(\frac{P^j}{h}\right) = T4aa + T4ab, \;\; T4aa = -\partial_i\left(\frac{P_j P^j}{h}\right), \;\; T4ab = \frac{P^j}{h}\partial_i P_j.$$

Since B is a gradient, we have $\partial_i B_j = \partial_j B_i$ and, therefore,

$$T4b = -T1b, \quad T2b = \partial_i B_j \frac{B^j}{h},$$

so that

$$T3b + T2b + T4ab = \frac{1}{2h}\partial_i(1 + D^2 + B_j B^j + P_j P^j) = \partial_i h = \partial_i(\frac{h^2}{h})$$

(by definition of h). Collecting all terms, we find

$$\begin{aligned}\partial_t P_i = T = T4aa + T4ab + T4b + T3a + T3b + T1a + T1b + T2a + T2b \\ = T4aa + T3a + T1a + T2a + \partial_i h \\ = -\partial_i\left(\frac{P_j P^j}{h}\right) - \partial_i\left(\frac{D^2}{h}\right) - \partial_j\left(\frac{P_i P^j}{h}\right) - \partial_j\left(\frac{B_i B^j}{h}\right) + \partial_i(\frac{h^2}{h}) \\ = \partial_i\left(\frac{1 + B_j B^j}{h}\right) - \partial_j\left(\frac{P_i P^j}{h}\right) - \partial_j\left(\frac{B_i B^j}{h}\right)\end{aligned}$$

(again by definition of h) and we have obtained the desired conservation law for P. Let us finally establish a conservation law for

$$h = h(D, B, P) = \sqrt{1 + D^2 + B_k B^k + P_k P^k}\,.$$

Notice that, from now on, it is forbidden to use $P = -DB$. Indeed, the conservation law should only follow from the augmented system and the property that B is a gradient. We get

$$\begin{aligned}\partial_t h = \frac{D\partial_t D + B^i \partial_t B_i + P^i \partial_t P_i}{h} \\ = \frac{D}{h}\partial_j\left(\frac{-DP^j + B^j}{h}\right) + \frac{B^i}{h}\partial_i\left(\frac{-B_j P^j + D}{h}\right) +\end{aligned}$$

$$+\frac{P^i}{h}\left(\partial_i\left(\frac{1+B_jB^j}{h}\right)-\partial_j\left(\frac{P_iP^j}{h}\right)-\partial_j\left(\frac{B_iB^j}{h}\right)\right)$$
$$= T1+T2+T3+T4+T5+T6+T7$$

where

$$T1=\frac{D}{h}\partial_j\left(\frac{-DP^j}{h}\right),\quad T2=\frac{D}{h}\partial_j\left(\frac{B^j}{h}\right)$$
$$T3=\frac{B^i}{h}\partial_i\left(\frac{-B_jP^j}{h}\right),\quad T4=\frac{B^i}{h}\partial_i\left(\frac{D}{h}\right)$$
$$T5=\frac{P^i}{h}\partial_i\left(\frac{1+B_jB^j}{h}\right)$$
$$T6=-\frac{P^i}{h}\partial_j\left(\frac{P_iP^j}{h}\right),\quad T7=-\frac{P^i}{h}\partial_j\left(\frac{B_iB^j}{h}\right)=-\frac{P^j}{h}\partial_i\left(\frac{B_jB^i}{h}\right)$$

We see that

$$T2+T4=\partial_i\left(\frac{DB^i}{h^2}\right),$$
$$T1+T6=-P^j\left(\frac{D}{h}\partial_j(\frac{D}{h})+\frac{P^i}{h}\partial_j(\frac{P_i}{h})\right)-\partial_jP^j\left(\frac{D^2+P^2}{h^2}\right)$$
$$=P^j\left(\frac{1}{h}\partial_j(\frac{1}{h})+\frac{B^i}{h}\partial_j(\frac{B_i}{h})\right)+\partial_jP^j\left(\frac{1+B_iB^i}{h^2}-1\right)$$

(by definition of h)

$$=\frac{P^j}{h}\partial_j(\frac{1}{h})+\frac{P^jB^i}{h^2}\partial_jB_i+\frac{P^jB_iB^i}{h}\partial_j(\frac{1}{h})+\partial_jP^j\left(\frac{1+B_iB^i}{h^2}-1\right),$$
$$=\frac{P^jB^i}{h^2}\partial_jB_i+\frac{P^j(1+B_iB^i)}{h}\partial_j(\frac{1}{h})+\partial_jP^j\left(\frac{1+B_iB^i}{h^2}\right)-\partial_jP^j$$
$$=\frac{P^jB^i}{h^2}\partial_jB_i+\partial_j\left(\frac{P^j(1+B_iB^i)}{h^2}\right)-T5-\partial_jP^j.$$

We also have

$$T3+T7=-\partial_i\left(\frac{P^jB_jB^i}{h^2}\right)-\frac{B^iP^j}{h^2}\partial_iB_j,$$

so that (since B is a gradient)

$$T1+T6+T3+T7+T5=\partial_j\left(\frac{P^j(1+B_iB^i)}{h^2}\right)$$
$$-\partial_i\left(\frac{P^jB_jB^i}{h^2}\right)-\partial_jP^j$$

and we have finally obtained

$$\partial_th=T1+T2+T3+T4+T5+T6+T7$$
$$=\partial_j\left(\frac{P^j(1+B_iB^i)}{h^2}\right)-\partial_i\left(\frac{P^jB_jB^i}{h^2}\right)+\partial_i\left(\frac{DB^i}{h^2}\right)-\partial_jP^j,$$

which is the desired conservation law and concludes the proof of Theorem 47.

9.1 Inhomogeneous incompressible Euler and Muskat equations

Another example when we can fruitfully derive degenerate parabolic equations out of entropic systems of conservation laws come from Fluid Mechanics. We start with the Euler equations, set on $\mathbb{T}^d$ for simplicity, of an incompressible *inhomogeneous* fluid subject to the action of an external potential Φ and we use the Boussinesq approximation:

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \nabla \cdot v = 0,$$

$$\overline{\rho}(\partial_t v + \nabla \cdot (v \otimes v)) + \nabla p = -\rho \nabla \Phi, \;\; \overline{\rho} = cst.$$

Notice that the density field ρ is advected by the velocity field v in the sense that

$$(\partial_t \rho + v \cdot \nabla)\rho = 0,$$

which is a consequence of *both* the continuity equation *and* the divergence-free condition on v.

Remark.
In geophysical Fluid Mechanics, the Boussinesq approximation, which is still widely used because its substantially simplifies numerical computations, amounts to neglecting the variation of the density in the acceleration term and substituting for it the constant $\overline{\rho}$ which should be considered as an average density (and, accordingly, ρ should be thought as the density minus its average rather than the density itself, which does not affect the equations since adding a constant to ρ does not modify them, thanks to the pressure term and the divergence-free condition). Anyway, this model is fully consistent with the least action principle without requiring any approximation, provided the action is defined by

$$\mathcal{A} = \int\int \left(\frac{1}{2}\overline{\rho}|v(t,x)|^2 - \rho(t,x)\Phi(x)\right) dxdt$$

subject to constraints:

$$\partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \nabla \cdot v = 0.$$

Indeed, introducing two Lagrange multipliers $\theta = \theta(t,x) \in \mathbb{R}$ and $q = q(t,x) \in \mathbb{R}$ for the constraints, we form the Lagrangian

$$\mathcal{L} = \int\int \left(\frac{1}{2}|v(t,x)|^2 - \rho(t,x)\Phi(x) - \partial_t\theta\rho - \nabla\theta \cdot \rho v - \nabla q \cdot v\right) dxdt$$

(where we have set $\overline{\rho} = 1$ for notational simplicity) and get, by successively varying v and ρ,

$$v = \rho\nabla\theta + \nabla q, \;\; \partial_t\theta + v \cdot \nabla\theta + \Phi = 0,$$

which leads back to

$$\partial_t v + \nabla \cdot (v \otimes v) + \nabla p = -\rho \nabla \Phi,$$

after elementary calculations, where p is related to q through:

$$p = \frac{1}{2}|v|^2 - v \cdot \nabla q.$$

[Strickly speaking this derivation is incomplete as $d > 3$ (which does not matter from a mechanical viewpoint) since the "Clebsch" decomposition $v = \rho \nabla \theta + \nabla q$ is too restrictive to describe a divergence-free vector field as $d > 3$. Then, additional Lagrange multipliers must be added in the action principle.]
End of remark.

$>$From now on, we simplify notations by setting $\overline{\rho} = 1$ and define the "Euler-Boussinesq" as

$$\text{EB}: \quad \partial_t v + \nabla \cdot (v \otimes v) + \nabla p = -\rho \nabla \Phi, \quad \partial_t \rho + \nabla \cdot (\rho v) = 0, \quad \nabla \cdot v = 0.$$

Observe the (formal) conservation of energy:

$$\frac{d}{dt} \int_{\mathbb{T}^d} \left(\frac{1}{2}|v(t,x)|^2 + \rho(t,x)\Phi(x) \right) dx = 0.$$

Also notice that for any suitable function Ψ we get the extra conservation

$$\frac{d}{dt} \int_{\mathbb{T}^d} \Psi(\rho(t,x)) dx = 0.$$

So, we may as well rewrite the conservation of energy as

$$\frac{d}{dt} \int_{\mathbb{T}^d} \{|v(t,x)|^2 + (\rho(t,x) + \Phi(x))^2\} dx = 0.$$

(just by taking $\Psi(r) = r^2$).

9.1.1 From Euler to Muskat by quadratic change of time

Let us again use the quadratic change of time method, applied to the Euler-Boussinesq (EB) system:

$$t \to \theta = t^2/2, \quad \text{new } \rho(\theta, x) = \text{old } \rho(\sqrt{2\theta}, x),$$

$$\text{new } v(\theta, x) = \frac{1}{\sqrt{2\theta}} \text{old } v(\sqrt{2\theta}, x),$$

After this change, the Euler-Boussinesq system becomes

$$\partial_\theta \rho + \nabla \cdot (\rho v) = 0, \quad \nabla \cdot v = 0,$$

$$v + 2\theta(\partial_\theta v + \nabla \cdot (v \otimes v)) + \nabla p = -\rho \nabla \Phi,$$

For small θ we just find, as asymptotic equations, the Muskat equations

$$\text{Mu} \quad \partial_\theta \rho + \nabla \cdot (\rho v) = 0, \quad \nabla \cdot v = 0, \quad v + \nabla p = -\rho \nabla \Phi.$$

9.1.2 Relative energy estimate for the Euler-Boussinesq equations

Proposition 49. *If (v, ρ) is a weak solution of Euler-Boussinesq with decreasing energy. Then, for all smooth fields $(\tilde{v}, \tilde{\rho})$ such that $\nabla \cdot \tilde{v} = 0$, we get the "relative energy" differential inequality*

$$\frac{d}{dt}\{||v-\tilde{v}||^2_{L^2(\mathbb{T}^d)} + ||\rho - \tilde{\rho}||^2_{L^2(\mathbb{T}^d)}\} \leq 2\int_{\mathbb{T}^d} \mathcal{L} + \mathcal{Q},$$

$$\mathcal{L} = (\tilde{v} - v)\cdot \tilde{E}_1 + (\tilde{\rho} - \rho)\tilde{E}_2$$

$$\mathcal{Q} = (\tilde{\rho} - \rho)(\tilde{v} - v)\cdot \nabla(\Phi + \tilde{\rho}) - (\tilde{v} - v)\otimes(\tilde{v} - v)\cdot(\nabla \tilde{v} + \nabla \tilde{v}^T),$$

$$\tilde{E}_1 = \partial_t \tilde{v} + \nabla \cdot (\tilde{v}\otimes \tilde{v}) + \tilde{\rho}\nabla\Phi, \;\; \tilde{E}_2 = \partial_t \rho + \partial_j(\rho v^j),$$

At this point, we have just Lions did for the homogeneous Euler equations in [174]. Then, still following Lions, we can get from the relative energy estimate a good concept of "dissipative" solutions to the Euler-Boussinesq system and get global existence and "weak-strong" stability (and uniqueness) results for the Euler-Boussinesq system.

9.1.3 "Dissipative solutions" for the Muskat system

>From the "relative energy" estimate obtained for the Euler-Boussinesq system, we almost immediately get a corresponding new concept of "dissipative solution" for the Muskat system just by using, again, the quadratic change of time method (as done, for example, for some geometric PDEs as in [71]. The result is a definition:

Definition 50. $(\rho, v) \in (C^0(L^2_w) \times L^2)([0,T]\times \mathbb{T}^d)$, *s.t.* $\nabla \cdot v = 0$, *is called a dissipative solution to the Muskat system if* $\forall (\tilde{\rho}, \tilde{v}) \in (W^{1,\infty} \times L^2)([0,T]\times\mathbb{T}^d)$ *s.t.* $\nabla\cdot\tilde{v} = 0$,

$$\forall t \in [0,T], \; \int_{\mathbb{T}^d} (\tilde{\rho} - \rho)(t,\cdot)^2 \leq e^{\,t\,\tilde{r}} \int_{\mathbb{T}^d} (\tilde{\rho} - \rho)(0,\cdot)^2$$

$-\int_0^t e^{(t-s)\tilde{r}} \int_{\mathbb{T}^d}\{2(v-\tilde{v})\cdot\tilde{E}_1 + 2(\rho - \tilde{\rho})\tilde{E}_2$

$$+|\tilde{v} - v|^2 + |\tilde{v} - v - (\tilde{\rho} - \rho)\nabla(\Phi + \tilde{\rho})|^2\}(s,x)dxds,$$

$\tilde{E}_1 = \tilde{v} + \tilde{\rho}\nabla\Phi,\; \tilde{E}_2 = \partial_t\tilde{\rho} + \tilde{v}\cdot\nabla\tilde{\rho},\; \tilde{r} = ||\nabla(\Phi + \tilde{\rho})||_{L^\infty}.$

10 Solution of initial value problems by convex minimization

Solving initial value problems by convex minimization is an old idea going back to the least square method for linear equations. For nonlinear systems of PDEs, in particular for parabolic equations and various gradient flows, there

has been many contributions, including Brezis-Ekeland, Ghoussoub, Mielke-Stefanelli, Visintin, etc... In a recent work [67], we have introduced a different approach, essentially based on the concept of weak, distributional solutions, that works for systems of hyperbolic conservation laws with a convex entropy, including the Euler equations of fluid mechanics, and the simple Burgers equation without viscosity.
More recently, we figured out how the method also applies to some parabolic problems, one of them being the quadratic porous medium equations. This case is so simple and the analysis is so straightforward that we have decided to describe it as our first example, although the strategy was first defined for the Euler equations of incompressible fluids.
In addition, let us mention that the convex optimization problems obtained by this method can be seen as some generalized variational mean-field games à la Lasry-Lions [89], with the peculiarity that they involve matrix-valued rather than scalar density fields, which is, to the best of our knowledge, still unusual in the theory of MFGs.

10.1 The elementary case of the porous media equations with quadratic non linearity

The porous media equations with quadratic non linearity (QPME, in brief), set on the periodic cube $\mathbb{T}^d$ (for simplicity), reads

$$\mathrm{QPME}: \quad \partial_t u = \Delta u^2/2, \quad u = u(t,x) \in \mathbb{R}, \quad t \geq 0, \quad x \in \mathbb{T}^d,$$

where u is, a priori, a nonnegative function that can be interpreted as a "density" function for some fluid moving in a porous medium.
N.B. From a statistical mechanics viewpoint, this equation, set on the entire euclidean space $\mathbb{R}^d$, can be obtained, as, more or less, in [176], as the macroscopic limit of the properly rescaled very simple (deterministic) system of N interacting particles:

$$\frac{dX_k}{dt} = \epsilon^{-1} \sum_{j=1,N} (X_k - X_j) \exp\left(-\frac{|X_k - X_j|^2}{\epsilon}\right),$$

$$u(t,x) \sim \frac{1}{N} \sum_{j=1,N} \delta(x - X_j(t)), \quad 1/N << \epsilon^d << 1.$$

This equation admits a Ljapunov (or "entropy") functional, namely

$$\int_{\mathbb{T}^d} u^2(t,x)dx,$$

for which we get, at least formally

$$\frac{d}{dt}\int_{\mathbb{T}^d} u^2(t,x)dx = -\int_{\mathbb{T}^d} u(t,x)|\nabla u|^2(t,x)dx,$$

We start with the rather absurd problem of minimizing, on a given finite time interval $[0, T]$, the time integral of the "entropy"

$$\int_Q u^2(t,x)dxdt, \quad Q = [0,T] \times \mathbb{T}^d,$$

among all weak (i.e. distributional) solutions in $L^2([0,T] \times \mathbb{T}^d)$ of the QPME

$$\partial_t u = \Delta u^2/2, \quad u = u(t,x) \geq 0, \quad t \geq 0, \quad x \in \mathbb{T}^d,$$

with a prescribed initial condition $u_0 \geq 0$, given, for simplicity, in $L^\infty(\mathbb{T}^d)$. A priori this problem is absurd since it is well known since the 80s that the Cauchy problem is uniquely solvable, for nonnegative distributional solutions, in $L^1(\mathbb{R}^d)$ [85], and that all L^p spaces (in particular L^2) are preserved by the corresponding semi-group of (nonnegative) solutions. Therefore, once u_0 is prescribed, there is a unique nonnegative admissible solution and the minimization problem looks trivial. However, we do not require that the weak solutions are nonnegative, which makes the problem more uncertain.
Anyway, this strange minimization problem admits a saddle point formulation which reads

$$I(u_0) = \inf_u \sup_\phi \int_Q \left(u^2 - 2\partial_t\phi u - \Delta\phi \, u^2 + 2u_0\partial_t\phi\right),$$

where the only constraints are:
i) for test function ϕ to be smooth and vanish at $t = T$;
ii) for function u to be square integrable on Q. By reversing the inf and the sup, we get a (non trivial!) relaxed problem

$$J(u_0) = \sup_\phi \inf_u \int_Q \left(u^2 - 2\partial_t\phi u - \Delta\phi \, u^2 + 2u_0\partial_t\phi\right).$$

At this level, we may just claim that $I(u_0) \geq J(u_0)$ and there may be a "duality gap" since the problem we started from is not formulated as a convex problem. The relaxed problem is very simple. Indeed, it is enough to perform the minimization in u pointwise in (t,x), since there is no more constraint on u:

$$J(u_0) = \sup_\phi \inf_u \int_Q \left(u^2 - 2\partial_t\phi u - \Delta\phi \, u^2 + 2u_0\partial_t\phi\right) =$$

$$\sup_\phi \int_Q \left(-\frac{(\partial_t\phi)^2}{1-\Delta\phi} + 2u_0\partial_t\phi\right), \quad \Delta\phi \leq 1, \quad \phi(T,\cdot) = 0.$$

Notice that the optimal value of u, for a given point (t,x), is given by

$$u = \frac{\partial_t\phi(t,x)}{1-\Delta\phi(t,x)},$$

under the condition that $\Delta\phi(t,x) < 1$ (otherwise the infimum in u is $-\infty$, unless both $\Delta\phi(t,x) = 1$ *and* $\partial_t\phi(t,x) = 0$ hold true simultaneously.).
Remark. Setting $q = \partial_t\phi$, $\sigma = 1 - \Delta\phi$, we get an alternative formulation:

$$J(u_0) = \sup_{\sigma,q} \int_Q \left(-\frac{q^2}{\sigma} + 2u_0\, q \right), \quad \partial_t\sigma + \Delta q = 0, \quad \sigma \geq 0, \quad \sigma(T,\cdot) = 1.$$

which is strongly reminiscent of some "optimal transport" problems (with quadratic cost), in there temporal (also known as Benamou-Brenier) formulation.

Analysis of the relaxed concave optimization problem

Let us now perform a rough analysis of our relaxed concave optimization problem, using what is already known about the QPME. To make our reasoning easier, we limit ourself to the easy case when u_0 is smooth and positive on $\mathbb{T}^d$. We want to prove

Theorem 51. *Any smooth positive solution $(t,x) \in Q = [0,T] \times \mathbb{T}^d \to u(t,x)$ of the quadratic porous medium equation QPME*

$$\partial_t u = \Delta u^2/2$$

can be recovered as

$$u = \frac{\partial_t\phi}{1 - \Delta\phi} ,$$

where ϕ solves the concave optimization problem

$$J(u_0) = \sup_{\phi} \int_Q \left(-\frac{(\partial_t\phi)^2}{1 - \Delta\phi} + 2u_0\partial_t\phi \right), \quad \Delta\phi \leq 1, \quad \phi(T,\cdot) = 0,$$

and satisfies

$$1 - \Delta\phi \geq (t/T)^{d/(d+2)}.$$

Proof. By standard parabolic regularity theory, the unique nonnegative weak solution $u(t,x)$ with smooth positive initial condition u_0 is a smooth and positive function of $(t,x) \in Q = [0,T] \times \mathbb{T}^d$. It is known [229] that all (nonnegative) solutions $u = u(t,x)$ of the QPME satisfy the Aronson-Bénilan estimate

$$\Delta u \geq -\kappa/t ,$$

where $\kappa = d/(d+2)$ just depends on d. Let us try to find a solution ϕ to the concave optimization problem just by solving the final value problem

$$\partial_t\phi = (1 - \Delta\phi)u, \quad \phi(T,\cdot) = 0,$$

i.e., in terms of $\alpha = 1 - \Delta\phi$,

$$\partial_t\alpha + \Delta(\alpha u) = 0, \quad \alpha(T,\cdot) = 1.$$

We claim that $\alpha(t,x) \geq (t/T)^\kappa$ follows from the Aronson-Bénilan estimate. Indeed, since u is smooth, we can write

$$\partial_t \alpha + \Delta(\alpha u) = \partial_t \alpha + u\Delta\alpha + 2\nabla\alpha \cdot \nabla u + \alpha\Delta u = 0$$

and, using both the maximum principle and the Aronson-Bénilan estimate, we get for $A(t) = \inf_{x\in\mathbb{T}^d} \alpha(t,x)$ the differential inequality

$$A'(t) \leq \kappa A(t)/t.$$

So, $\log A(T) - \log A(t) \leq \kappa(\log T - \log t)$, and therefore $A(t) \geq (t/T)^\kappa$ (since $A(T) = 1$). This estimate shows that the function $\alpha = 1 - \Delta\phi$ stays positive on $]0,T] \times \mathbb{T}^d$. Let us now finally show that ϕ is optimal for the concave maximization problem. For that purpose, let us just evaluate

$$j = \int_Q \left(-\frac{(\partial_t\phi)^2}{1-\Delta\phi} + 2u_0\partial_t\phi \right).$$

which, by definition of $J(u_0)$, is certainly bounded from above by $J(u_0)$. Since u solves the QPME with initial condition u_0, we have

$$\int_Q \left(2\partial_t\phi u + \Delta\phi u^2 - 2\partial_t\phi u_0\right) = 0.$$

Thus, since ϕ solves $\partial_t\phi = (1-\Delta\phi)u$,

$$j = \int_Q \left(-\frac{(\partial_t\phi)^2}{1-\Delta\phi} + 2u\partial_t\phi + \Delta\phi u^2 \right) = \int_Q u^2$$

which shows that ϕ is optimal since, by construction,

$$J(u_0) \geq j = \int_Q u^2 \geq I(u_0) \geq J(u_0).$$

End of Proof.

Various comments

1) Through additional technical work, this proof should extend to *all* initial conditions in $L^2(\mathbb{R}^d)$. The theory should also apply to the case of the entire euclidean space $\mathbb{R}^d$ and to the famous "Barenblat profiles", that have compact support and saturate the Aronson-Bénilan estimate [229].
2) Notice that, strictly speaking, we have not shown the uniqueness of a maximizer for the concave maximization problem.
3) Our formulation in terms of convex optimization might be a useful way of getting new regularity results for the QPME. This problem is of current interest since new regularity results have been obtained:
a) in [142] by Gess, Sauer and Tadmor, for the porous medium equation, through quite unusual methods in the elliptic setting such as "average lemmas" coming from kinetic theory [151];
b) in [150] by Goldman and Otto, for the quadratic optimal transport problem in its temporal "Benamou-Brenier" formulation, which looks very similar to the relaxed concave optimization problaim we have just obtained for the QPME.

10.2 A much more challenging parabolic example: the quantum diffusion equation (QDE)

Just to indicate, without any further analysis, a highly non trivial example of a parabolic system for which the initial value problem could be fruitfully addressed in terms of convex optimization, let us mention the so-called quantum diffusion equation (written as a system in weak form according to [145] sect. 1.8):

$$\text{QDE}: \qquad \partial_t u + \Delta^2 u - D^2 : \frac{g \otimes g}{u} = 0, \quad g = \nabla u,$$

where $u : (t,x) \in Q = [0,T] \times \mathbb{T}^d \to u(t,x) \geq 0$, for which

$$\int_{\mathbb{T}^d} \frac{|g(t,x)|^2}{2u(t,x)} dx$$

is a Ljapunov function, or an "entropy" in Otto's framework of gradient flows for transportation metrics [145].
We start by minimimizing the time integral over $[0,T]$ of the entropy among all weak solutions of QDE with given initial condition u_0, which leads to the saddle point problem:

$$I(u_0) = \inf_{(u \geq 0,\ g)} \quad \sup_{(\phi,P)} - \int_{\mathbb{T}^d} u_0(x)\phi(0,x)dx$$

$$+ \int_Q \left(\frac{|g|^2}{2u} - \partial_t \phi u + \Delta^2 \phi u - D^2 \phi : \frac{g \otimes g}{u} - P \cdot g - u \nabla \cdot P \right)(t,x)dxdt,$$

(where $P = P(t,x) \in \mathbb{R}^d$ is a Lagrange multiplier for constraint $g = \nabla u$). Reversing the inf and the sup leads to the desired relaxed concave maximization problem. By minimizing in g (pointwise in (t,x) since there is no constraint on g), we first get

$$J(u_0) = \sup_{(\phi,P)} \quad \inf_{u \geq 0} - \int_{\mathbb{T}^d} u_0(x)\phi(0,x)dx$$

$$+ \int_Q u(t,x) \left(-\frac{1}{2}(\mathbb{I}_d - 2D^2\phi)^{-1} : P \otimes P - \partial_t \phi + \Delta^2 \phi - \nabla \cdot P \right)(t,x)dxdt,$$

where $\mathbb{I}_d$ is the identity matrix and $\phi : (t,x) \in Q = [0,T] \times \mathbb{T}^d \to \phi(t,x) \in \mathbb{R}$ is subject to $D^2\phi \leq \mathbb{I}_d$ and $\phi(T,\cdot) = 0$. Then, after minimizing, again pointwise, in $u \geq 0$, we finally obtain:

$$J(u_0) = \sup_{(\phi,P)} \quad - \int_{\mathbb{T}^d} u_0(x)\phi(0,x)dx,$$

where, ϕ is subject, again, to $D^2\phi \leq \mathbb{I}_d$ and $\phi(T,\cdot) = 0$ and also to the pointwise inequality:

$$\partial_t \phi - \Delta^2 \phi + 1/2(I_d - 2D^2\phi)^{-1} : (P \otimes P) + \nabla \cdot P \leq 0,$$

for some unknown vector field $P : (t,x) \in [0,T] \times \mathbb{T}^d \to P(t,x) \in \mathbb{R}^d$.

10.3 The class of "entropic conservation laws"

They read

$$\partial_t U + \nabla \cdot (F(U)) = 0, \;\; U = U(t,x) \in \mathcal{W} \subset \mathbb{R}^m, \;\; x \in \mathbb{T}^d,$$

where F (usually called the "flux function") enjoys the symmetry property

$$\sum_{\beta=1}^{m} \partial_\beta \mathcal{E}(W) \partial_\alpha F^{i\beta}(W) = \partial_\alpha Q^i(W), \; \forall W \in \mathcal{W},$$

for some pair of functions $(\mathcal{E}, Q) : \mathcal{W} \to \mathbb{R}^{1+d}$, where $\mathcal{W}$ is an open convex subset of $\mathbb{R}^m$ and $\mathcal{E}$ (usually called "entropy") is strictly convex over W. This stuctural condition implies that, whenever $U = U(t,x)$ is a smooth solution of the system, we get the additional conservation law

$$\partial_t(\mathcal{E}(U)) + \nabla \cdot (Q(U)) = 0.$$

Of course the simplest example is the so-called "inviscid Burgers" equation, where $U = u(t,x)$ is a real-valued function of a single space variable x with the simplest nonlinear flux function $F = u^2/2$:

$$\partial_t u + \partial_x (u^2/2) = 0.$$

It is well established that, in most situations, such systems admit smooth solutions that blow up (in Lipschitz norm) after a finite time, phenomenon known as "shock formation", by reference to compressible gas dynamics.

A canonical example: the Euler equations of isothermal compressible fluids.

They simply read

$$\partial_t \rho + \nabla \cdot q = 0, \;\; \partial_t q + \nabla \cdot (\frac{q \otimes q}{\rho}) + \nabla \rho = 0,$$

and fit into the general framework just by defining

$$U = (\rho, q) \in \mathcal{W} =]0, +\infty[\times \mathbb{R}^3, \; F = (q, \frac{q \otimes q}{\rho} + I_3 \rho), \; \mathcal{E} = -\frac{|q|^2}{2\rho} - \rho \log \rho$$

The least square approach?

Given U_0 on $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ and $T > 0$, if $F(U)$ is linear in U, the least square method obviously leads to a (degenerate) convex problem

$$\inf_{U(t=0,\cdot)=U_0} \int_{[0,T]\times\mathbb{T}^d} |\partial_t U + \nabla \cdot (F(U))|^2$$

but this is no longer true for nonlinear systems.

We are going to use another idea based on weak solutions, just as we did for the quadratic porous medium equation (QPME).

Minimization approach to the Cauchy problem

Given U_0 on $D = \mathbb{R}^d/\mathbb{Z}^d$ and $T > 0$, we minimize the entropy among all weak solutions U of the Cauchy pb:

$$I(U_0) = \inf_U \int_0^T \int_D \mathcal{E}(U), \quad U = U(t,x) \in \mathcal{W} \subset \mathbb{R}^m \text{ subject to}$$

$$\int_0^T \int_D \partial_t A \cdot U + \nabla A \cdot F(U) + \int_D A(0,\cdot) \cdot U_0 = 0$$

for all smooth $A = A(t,x) \in \mathbb{R}^m$ with $A(T,\cdot) = 0$. The problem is not trivial since there may be many weak solutions starting from U_0 which are not entropy-preserving (by "convex integration" à la De Lellis-Székelyhidi) [107, 108, 109, **?**]. We get the resulting saddle-point problem

$$\inf_U \ \sup_A \ \int_0^T \int_D \mathcal{E}(U) - \partial_t A \cdot U - \nabla A \cdot F(U)$$
$$- \int_D A(0,\cdot) \cdot U_0$$

where $A = A(t,x) \in \mathbb{R}^m$ is smooth with $A(T,\cdot) = 0$.
Here U_0 is the initial condition and T the final time.
N.B. The supremum in A exactly encodes that U is a weak solution with initial condition U_0, all test functions A acting like Lagrange multipliers.

Reversing infimum and supremum...

This leads to a *concave* maximization problem in A, namely

$$J(U_0) = \sup_{A(T,\cdot)=0} \ \inf_U \ \int_0^T \int_D \mathcal{E}(U) - \partial_t A \cdot U - \nabla A \cdot F(U) - \int_D A(0,\cdot) \cdot U_0$$
$$= \sup_{A(T,\cdot)=0} \ \int_0^T \int_D -G(\partial_t A, \nabla A) - \int_D A(0,\cdot) \cdot U_0$$

where G is defined by

$$G(E,B) = \sup_{V \in \mathcal{W} \subset \mathbb{R}^m} E \cdot V + B \cdot F(V) - \mathcal{E}(V), \ (E,B) \in \mathbb{R}^m \times \mathbb{R}^{m \times d}.$$

Notice that G is automatically convex (but presumably degenerate!). Thus we have obtained a (possibly degenerate) space-time elliptic system in A, which is reminiscent of those appearing in optimal transport theory (as will be discussed later on). Here is the paradox! How a convex optimization problem could be compatible with a well-posed evolution problem? For instance, if G were just a square, we would get

$$\sup_A \int_0^T \int_D -|\partial_t A|^2 - |\nabla A|^2 - \int_D A(0,\cdot) \cdot U_0$$

which would correspond to an ill-posed equation for A:

$$\partial_{tt}^2 A + \Delta A = 0.$$

The answer to the paradox is that, in our construction, G is very likely to be convex *degenerate* which is presumably still compatible with the solution of a well-posed initial value problem.

Main results for entropic conservation laws

Theorem 52. *If U is a smooth solution to the IVP and T is not too large, in the sense that*

$$\forall\, t, x, V \in \mathcal{W}, \mathcal{E}"(V) - (T-t)F"(V) \cdot \nabla(\mathcal{E}'(U(t,x))) > 0,$$

then U can be recovered from the concave maximization problem which admits $A(t,x) = (t-T)\mathcal{E}'(U(t,x))$ as solution.

Notice that the smallness condition requires, in particular,

$$\mathcal{E}"(V) - TF"(V) \cdot \nabla(\mathcal{E}'(U_0(x))) > 0,$$

and definitely restricts the choice of T with respect to U_0. This is clearly a drawback of the theory. So we should worry about the generic apparition of shock waves and give up any hope to be able to solve the initial value problem for arbitrarily large values of T. However, in the very elementary case of the Burgers equation, all entropy solutions (in the sense of Kruzhkov, see [106] for this concept of solutions) can be recovered, for arbitrarily large T, but in some unusual way. More precisely

Theorem 53. *If u is a Kruzhkov solution of the inviscid Burgers equation on some fixed time interval T with initial condition u_0, then the relaxed convex optimisation problem enables us to recover not necessarily the Kruzhkov solution itself but rather the unique solution $u^T(t,x)$ of the inviscid Burgers equation enjoying the following properties:*
1) u^T and u coincide at the final time T;
2) u^T is shock free up to time $t = T$ (not included).
In general, the initial value of u^T differs from u_0, unless no shock have formed before T.

So, our method is able to recover the right Kruzhkov entropy but only at the final given time T, as soon as shock have formed before T. This result is also a new answer to the paradox discussed earlier. Something is left from the degenerate space-time ellipticity of the convex minimization problem in the sense that the smoothest possible solution of the inviscid Burgers equation compatible with the right final solution is selected, just by substituting for the given initial condition u_0 another one, namely $u^T(0,\cdot)$.

10.3.1 Examples and matrix-valued variational mean-field games

Before moving to the proofs of these results, let us look more carefully at explicit examples of hyperbolic conservation laws, such as the Burgers equation (without viscosity) and the much more challenging Euler equations. In the elementary example of the Burgers equation, the maximization problem in A simply reads

$$\sup_A \int_{[0,T]\times\mathbb{T}} -\frac{(\partial_t A)^2}{2(1-\partial_x A)} - \int_{\mathbb{T}} A(0,\cdot)u_0.$$

with $A = A(t,x) \in \mathbb{R}$ subject to $A(T,\cdot) = 0$, $\partial_x A \leq 1$. Introducing

$$\rho = 1 - \partial_x A \geq 0, \;\; q = \partial_t A,$$

we get:

$$\sup_{(\rho,q)} \{ \int_{[0,T]\times\mathbb{T}} -\frac{q^2}{2\rho} - qu_0 \mid \partial_t\rho + \partial_x q = 0, \;\; \rho(T,\cdot) = 1 \}.$$

N.B. This problem can be interpreted, in our opinion, as the "ballistic" version (à la Ghoussoub) of the optimal transport problem with quadratic cost and, as well, as a rather trivial example of MFG à la Lasry-Lions (without noise nor interaction) of variational type. So we may expect more interesting connections with MFG, when addressing more complex equations than the inviscid Burgers one!

In addition, the resulting problem

$$\sup_{(\rho,q)} \{ \int_{[0,T]\times\mathbb{T}} -\frac{q^2}{2\rho} - qu_0 \mid \partial_t\rho + \partial_x q = 0, \;\; \rho(T,\cdot) = 1 \}$$

is so close to an optimal transport problem (in its so-called Benamou-Brenier formulation) that, at the computational level, it differs from it just by two lines of (fortran) code, when using the algorithm designed in [23].

Let us now move to the more sophisticated case of the isothermal Euler equations:

$$\partial_t\rho + \nabla\cdot q = 0, \;\; \partial_t q + \nabla\cdot(\frac{q\otimes q}{\rho}) + \nabla\rho = 0.$$

In this case, we are going to see that our convex optimization problem to solve the IVP can be interpreted as a *generalized* (variational deterministic) mean-field game involving fields of nonnegative symmetric matrices instead of density fields. Indeed, we get the convex optimization problem

$$\int_{[0,T]\times D} \exp(u)\exp(\frac{1}{2}Q\cdot M^{-1}\cdot Q) + \int_D \sigma_0\rho_0 + w_0\cdot q_0,$$

among all fields $u = u(t,x) \in \mathbb{R}$, $Q = Q(t,x) \in \mathbb{R}^d$, $M = M(t,x) = M^t(t,x) \in \mathbb{R}^{d\times d}$, $\quad M \geq 0, \quad$ of form:

$$u = \partial_t\sigma + \partial^i w_i, \; Q_i = \partial_t w_i + \partial_i\sigma, \; M_{ij} = \delta_{ij} - \partial_i w_j - \partial_j w_i,$$

where σ and w must vanish at $t = T$.
Finally, let us discuss the Euler equations of incompressible fluids

$$\partial_t q + \nabla \cdot (q \otimes q) = -\nabla p, \quad \nabla \cdot q = 0,$$

where q is prescribed at $t = 0$ and p is now a Lagrange multiplier for constraint $\nabla \cdot q = 0$. We get again a generalized MFG for measures valued in the cone of semi-definite symmetric matrices.

$$\sup_{(M,Q)} \quad -\int_{[0,T]\times D} q_0 \cdot Q + \frac{1}{2}\, Q \cdot M^{-1} \cdot Q,$$

where now Q is a vector field (not necessarily divergence-free) and $M = M^t \geq 0$ is a field of semi-definite symmetric matrices subject to

$$M_{ij}(T, \cdot) = \delta_{ij}, \quad \partial_t M_{ij} = \partial_j Q_i + \partial_i Q_j + 2\partial_i \partial_j (-\Delta)^{-1} \partial_k Q^k.$$

Inviscid Burgers equation : $\partial_t u + \partial_x(u^2/2) = 0$, $u = u(t, x)$, $x \in \mathbb{R}/\mathbb{Z}$, $t \geq 0$.
Formation of two shock waves. (Vertical axis: $t \in [0, 1/4]$, horizontal axis: $x \in \mathbb{T}$.)

11 A dissipative least action principle for approximations of the Euler equations

Our purpose is to exhibit relevant approximations of the Euler equations for which a *modified* least action principle can be designed that can include *energy dissipation.* There are examples, typically in infinite dimension (but not necessarily), of *formally* hamiltonian systems which *do not* necessarily *preserve* the energy because of some hidden dissipative mechanism:
i) the (inviscid) Burgers equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(\frac{u^2}{2}) = 0, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R} \to u(t, x) \in \mathbb{R};$$

ii) the Euler equations of incompressible fluids: at least at the physical level, it is often believed that the energy could dissipate according to Kolmogorov's "K41" theory of turbulence [137].

Let us start the discussion with special examples of finite dimensional dynamical systems for which a dissipative version of the least action principle can be designed.

11.1 Finite dimensional examples

Given an Euclidean space H (or more generally a Hilbert space) with norm $||\cdot||$ and a potential $Q : H \to \mathbb{R}$,

$$\frac{1}{2}||V_t||^2 + Q[X_t]$$

is the conserved energy (or Hamiltonian) for the dynamical system

$$\frac{dV_t}{dt} = -\nabla Q[X_t], \quad \frac{dX_t}{dt} = V_t, \quad (X_t, V_t) \in H \times H.$$

As well known, its solutions can be obtained from the "least action principle" by looking for critical points of the "action"

$$\int_{t_0}^{t_1} \frac{1}{2}||\frac{dX_t}{dt}||^2 - Q[X_t] \quad dt,$$

among all curves $t \in [t_0, t_1] \to X_t$ with fixed values at t_0 and t_1.
We are going to define a special class of hamiltonian systems (in finite dimension), for which a *modified* least action principle can be designed that can

include *energy dissipation*. This issue has been already discussed by various authors, Shnirelman and Wolansky, for instance [218, 233]. The systems we are going to discuss are very special but, among them, we will get discrete or approximate versions of the Euler model of incompressible fluids.

Let H be a Euclidean space and S a bounded closed subset. Set

$$Q[X] = -\frac{1}{2}\mathrm{dist}^2(X,S) = -\inf_{s\in S}\frac{||X-s||^2}{2}$$

and consider the corresponding dynamical system

$$\frac{d^2X_t}{dt^2} = -\nabla Q[X_t]$$

N.B.: Q is semi-convex, but not smooth (unless S is convex).
Indeed: $Q[X] = -\frac{1}{2}||X||^2 + R[X]$, where $R[X] = \sup_{s\in S}((X,s)) - \frac{1}{2}||s||^2$ is convex.

11.2 The main example and the Vlasov-Monge-Ampère system

Let us now describe our main example. Let $\{A(1),\cdots,A(N)\}$ be a cubic lattice of N points approximating $D = [-1/2,1/2]^d \subset \mathbb{R}^d$ as N tends to infinity. Define

$$H = (\mathbb{R}^d)^N, \quad S = \{(A(\sigma_1),\cdots,A(\sigma_N)) \in H, \quad \sigma \in \mathcal{S}_n\}$$

(where $\mathcal{S}_N$ denotes the group of all permutations of the first N integers, while $|\cdot|$ and $||\cdot|| =$ are the euclidean norms respectively on $\mathbb{R}^d$ and $\mathbb{R}^{Nd}$.)
Then, the dynamical system introduced in the previous subsection reads, after elementary calculations,

$$\beta\frac{d^2X_t(\alpha)}{dt^2} = X_t(\alpha) - A(\sigma_{opt}(\alpha)) \ , \quad X_t(\alpha)\in\mathbb{R}^d, \ \ \alpha = 1,\cdots,N \qquad (35)$$

$$\sigma_{opt} = \mathrm{Arginf}\ {}_{\sigma\in\mathcal{S}_N} \quad \sum_{\alpha=1}^{N} |X_t(\alpha) - A(\sigma(\alpha))|^2 \qquad (36)$$

with $\beta = 1$, involving, at each time t, a discrete optimal transport problem.
This system was introduced, in the case $\beta = -1$, in [50], where its hydrodynamic limit to the Euler equations has been established.
Notice that, as $d = 1$, this system reduces to

$$\beta\frac{d^2X_t(\alpha)}{dt^2} = X_t(\alpha) - \frac{1}{2N}\sum_{\alpha'\neq\alpha}\mathrm{sgn}(X_t(\alpha) - X_t(\alpha')).$$

This describes the Newtonian gravitational interaction of N parallel planes as $\beta = 1$ (with a global neutralization of the total mass, expressed by the linear term X_t).
The continuous version, involving the Monge-Ampère equation, closely related to optimal transport theory, was introduced by B. and Loeper [77], and studied by Cullen, Gangbo, Pisante [104], Ambrosio-Gangbo [7]. We find

$$\partial_t f(t,x,\xi) + \nabla_x \cdot (\xi\ f(t,x,\xi)) - \nabla_\xi \cdot (\nabla_x \varphi(t,x) f(t,x,\xi)) = 0 \tag{37}$$

$$\det(\mathbb{I} - \beta D_x^2 \varphi(t,x)) = \int_{\mathbb{R}^d} f(t,x,\xi) d\xi, \quad (t,x,\xi) \in \mathbb{R} \times D \times \mathbb{R}^d. \tag{38}$$

This fully nonlinear version of the Vlasov-Poisson system is related to Electrodynamics ($\beta = -1$) and Gravitation ($\beta = 1$). The formal limit $\beta = 0$ reads

$$\partial_t f + \nabla_x \cdot (\xi\ f) - \nabla_\xi \cdot (\nabla_x p\ f) = 0, \int_{\mathbb{R}^d} f(t,x,\xi) d\xi = 1,$$

where $p = p(t,x)$ substitutes for φ as a Lagrange multiplier of constraint $\int f d\xi = 1$. It can be understood as a "kinetic formulation" of the Euler equations of homogeneous incompressible fluids (see [43, 48], for this concept). Classical solutions (v,p) to the Euler equations correspond to very special and singular solutions of the kinetic version of form

$$f(t,x,\xi) = \delta(\xi - v(t,x)).$$

11.3 Conservative solutions à la Bouchut-Ambrosio

Let us go back to the general case, where H and S can be chosen freely, respectively as an Euclidean space and a bounded closed subset. The dynamical system

$$\frac{d^2 X_t}{dt^2} = -\nabla Q[X_t]$$

with $Q[X] = -\frac{1}{2}||X||^2 + R[X]$, where $R[X] = \sup_{s \in S}((X,s)) - \frac{1}{2}||s||^2$ is convex, Lipschitz continuous, but not smooth (unless S is convex), cannot be treated by the usual Cauchy-Lipschitz theory. However the second derivatives of R are nonnegative bounded measures and we may apply the DiPerna-Lions theory [117], as generalized by Bouchut and Ambrosio to second-order ODEs with "coefficients of bounded variation" [3, 35]: for "almost every initial condition"

$$(X_0,\ \frac{dX_0}{dt}) \in H \times H,$$

$$\frac{d^2 X_t}{dt^2} = -\nabla Q[X_t] = X_t - \nabla R[X_t]$$

admits a global $C^{1,1}$ solution, unique in a sense precised by Ambrosio.
Such a solution is "conservative" and time-reversible. For the system of particles discussed in the previous subsection, in particular in the framework of 1D-Newtonian gravitation, this corresponds to elastic, non-dissipative collisions.

11.4 Rewriting of the action for "good" curves

There is a subset $N \subset H$, which is small in both the Baire category sense and the Lebesgue measure sense (but not empty unless S is convex), outside of which every point $X \in H \setminus N$ admits a unique closest point $\pi[X]$ on S and

$$Q = -\frac{1}{2}\text{dist}^2(\cdot, S)$$

is differentiable at X with:

$$-\nabla Q[X] = X - \pi[X], \quad Q[X] = -\frac{1}{2}||X - \pi[X]||^2 = -\frac{1}{2}||\nabla Q[X]||^2.$$

So, the potential can be rewritten as a negative squared gradient.
Thus, for any "good" curve which almost never hits the bad set N, the action can be written

$$\frac{1}{2}\int_{t_0}^{t_1} ||\frac{dX_t}{dt}||^2 + ||\nabla Q[X_t]||^2 \quad dt$$

which can be rearranged as a perfect square up to a boundary term that does not play any role in the least action principle

$$\frac{1}{2}\int_{t_0}^{t_1} ||\frac{dX_t}{dt} + \nabla Q[X_t]||^2 \quad dt \quad - Q[X_{t_1}] + Q[X_{t_0}].$$

11.5 Gradient-flow solutions as special least-action solutions

Due to the very special structure of the action, we find as particular least action solutions any solution to the first-order "gradient-flow equation"

$$\frac{dX_t}{dt} = -\nabla Q[X_t]$$

(somewhat like "instantons" in Yang-Mills theory). However, this is correct only when $t \to X_t \in H$ is a "good" curve (i.e. almost never hits the "bad set" where Q is not differentiable).

11.6 Global dissipative solutions of the gradient-flow

Since Q is semi-convex, we may use the classical theory of maximal monotone operators (going back to the 70', as in the book by H. Brezis [83]) to solve the initial value problem for the gradient-flow equation.
For each initial condition, there is a unique global solution s.t

$$\frac{d_+X_t}{dt} = -\overline{\nabla} Q[X_t] \ , \ \ \forall t \geq 0., \quad X \in C^0([0, +\infty[, H). \tag{39}$$

Here, $\frac{d_+}{dt}$ denotes the right-derivative at t, and, for each X,

$$\overline{\nabla}Q[X] = -X + \overline{\nabla}R[X]$$

where $\overline{\nabla}R[X]$ is the "relaxed" gradient of the convex function R at point X, i.e. the unique $w \in H$ with lowest norm, $||w||$, such that

$$R[Z] \geq R[X] + ((w, Z - X)), \quad \forall Z \in H.$$

The relaxed gradient is well defined for every X and extends the usual gradient to the "bad set" N. These solutions in the sense of maximal monotone operator theory are in general not conservative solutions (in the sense of Bouchut-Ambrosio) to the original dynamical system. Indeed, they allow velocity jumps and are generally only Lipshitz continuous and not C^1.
However, they have interesting dissipative features. Indeed, the velocity may jump with an instantaneous loss of kinetic energy.
In the case of one-dimensional gravitating particles, these jumps precisely correspond to sticky collisions [75, 74]. The bad set N is just the collision set and the relaxed gradient precisely encodes sticky collisions instead of elastic collisions.

11.7 A proposal for a modified action

The conservative solutions, that are only defined for almost every initial condition, manage to hit the bad set only for a negligible amount of time, while the gradient flow solutions enjoy very much staying in it as soon as they enter it.
Our proposal is to pick up the nice dissipative property of the gradient flow solutions and to lift them to the full dynamical system. For that purpose, we introduce the "modified action"

$$\int_{t_0}^{t_1} ||\frac{dX_t}{dt}||^2 + ||\overline{\nabla}Q[X_t]||^2 \quad dt \tag{40}$$

which favors "bad" curves that stay on the "bad set" for a while. Let us recall that $\overline{\nabla}Q$ denotes the "relaxed" gradient of the semi-convex function

$$Q[X] = -\frac{1}{2}\mathrm{dist}^2(X, S) = -\frac{1}{2}||X||^2 + \sup_{s \in S}\{((X, s)) - \frac{1}{2}||s||^2\}. \tag{41}$$

12 Stochastic and quantum origin of the dissipative least action principle

Using large deviation principles (or alternatively the concept of guiding wave coming from quantum mechanics), we will derive, following [66] and from essentially nothing (namely N independent Brownian particles without any interaction nor external potential), the dissipative least action principle (40,41), for the special system (35,36), in the "gravitational" case $\beta = 1$. Let us recall that this system is a discretization of the Vlasov-Monge-Ampère system (37,38) as well as an approximation of the Euler equations.
The first step of our analysis is very much related to the Schrödinger problem, as analyzed by Christian Léonard [170] and also to recent results by Robert Berman on permanental processes related to Kählerian Geometry [**?**].

12.1 Localization of a Brownian point cloud

Given a point cloud

$$\{A(\alpha) \in \mathbb{R}^d,\ \alpha = 1, \cdots, N\},$$

we consider N independent Brownian curves issued from this cloud

$$Y_t(\alpha) = A(\alpha) + \sqrt{\epsilon} B_t(\alpha), \quad \alpha = 1, \cdots, N.$$

At a fixed time $T > 0$, the probability for the moving cloud to reach position $X = (X(\alpha),\ \alpha = 1, \cdots, N) \in \mathbb{R}^{dN}$ has density

$$\frac{1}{Z} \sum_{\sigma \in \mathcal{S}_N} \prod_{\alpha=1}^{N} \exp\left(-\frac{|X(\alpha) - A(\sigma(\alpha))|^2}{2\epsilon T}\right)$$

$$= \frac{1}{Z} \sum_{\sigma \in \mathcal{S}_N} \exp\left(-\frac{||X - A_\sigma||^2}{2\epsilon T}\right)$$

(here $\mathcal{S}_N$ denotes the group of all permutations of the first N integers, while $|\cdot|$ and $||\cdot|| =$ are the euclidean norms respectively on $\mathbb{R}^d$ and $\mathbb{R}^{Nd}$.)
Since

$$-\epsilon \log \frac{1}{Z} \sum_{\sigma \in \mathcal{S}_N} \exp\left(-\frac{||X - A_\sigma||^2}{2\epsilon T}\right) \sim \frac{1}{2T} \inf_{\ \sigma \in \mathcal{S}_N} ||X - A_\sigma||^2$$

as $\epsilon \to 0$, an observer at time T feels that the particles arrived at $X_T \in \mathbb{R}^{dN}$, have travelled along straight lines by "optimal transport"

$$X_t = \left(1 - \frac{t}{T}\right) A_{\sigma_{opt}} + \frac{t}{T} X_T\ , \quad \sigma_{opt} = \mathrm{Arginf}_{\ \sigma \in \mathcal{S}_N}\ ||X_T - A_\sigma||^2.$$

This formula implies (through a simple argument)

$$\frac{dX_t}{dt} = \frac{X_t - A_{\sigma_{opt}}}{t}\ , \quad \sigma_{opt} = \text{Arginf}\ _{\sigma \in \mathcal{S}_N}\ \ ||X_t - A_\sigma||^2.$$

The resulting "deterministic" process is, as a matter fact, just the output of the pure observation of a random process as the level of noise vanishes. >From a physical viewpoint, it is equivalent to the Zeldovich model in Cosmology [237, 217, 138, 72]

12.2 An alternative viewpoint: the pilot wave

Introducing the heat equation in the space of "clouds" $X \in \mathbb{R}^{Nd}$

$$\frac{\partial \rho}{\partial t}(t, X) = \frac{\epsilon}{2} \triangle \rho(t, X), \quad \rho(t = 0, X) = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \delta(X - A_\sigma),$$

we follow the "pilot wave" à la de Broglie, solving the ODE $\frac{dX_t}{dt} = v(t, X_t)$ with "velocity" $v(t, X) = -\frac{\epsilon}{2} \nabla_X \log \rho(t, X)$ and find

$$\frac{dX_t}{dt} = \frac{X_t - < A >}{2t}\ , \ < A >= \frac{\sum_{\sigma \in \mathcal{S}_N} A_\sigma \exp(\frac{-||X_t - A_\sigma||^2}{2\epsilon t})}{\sum_{\sigma \in \mathcal{S}_N} \exp(\frac{-||X_t - A_\sigma||^2}{2\epsilon t})}.$$

[As a matter of fact, a similar calculation also works for the free bosonic Schrödinger equation: $(i\partial_t + 1/2 \triangle)\psi = 0, \quad \psi(0, X) = \sum_\sigma \exp(-||X - A_\sigma||^2/a^2), \quad v = \mathcal{I}m \nabla \log \psi$.]

Using exponential time $t = \exp(2\theta)$, we may also write:

$$\frac{dX_\theta}{d\theta} = X_\theta - < A >\ , \quad < A > = \frac{\sum_{\sigma \in \mathcal{S}_N} A_\sigma \exp(\frac{-||X_\theta - A_\sigma||^2}{2\epsilon \exp(2\theta)})}{\sum_{\sigma \in \mathcal{S}_N} \exp(\frac{-||X_t - A_\sigma||^2}{2\epsilon \exp(2\theta)})}.$$

As $\epsilon \to 0$, we obtain (39) in the sense of maximal monotone operator theory:

$$\frac{d_+ X_\theta}{d\theta} = -\overline{\nabla} Q[X_\theta]\ , \quad Q[X] = \inf\ _{\sigma \in \mathcal{S}_N}\ \ ||X - A_\sigma||^2/2.$$

12.3 Large deviations of the pilot system

Let us add some noise η to the "guided" trajectories (with fixed ϵ)

$$\frac{dX^\epsilon_\theta}{d\theta} = X^\epsilon_\theta - < A > + \eta \frac{dB_\theta}{d\theta}\ , \quad < A > = \frac{\sum_{\sigma \in \mathcal{S}_N} A_\sigma \exp(\frac{-||X^\epsilon_\theta - A_\sigma||^2}{2\epsilon \exp(2\theta)})}{\sum_{\sigma \in \mathcal{S}_N} \exp(\frac{-||X^\epsilon_\theta - A_\sigma||^2}{2\epsilon \exp(2\theta)})}$$

For ϵ fixed and $\eta \to 0$, we first get the corresponding large deviation rate function. Then, as $\epsilon \to 0$ we can pass to the $\Gamma-$limit (*) and obtain the dissipative action (40,41), namely

$$\int (||\frac{dX_\theta}{d\theta}||^2 + ||\overline{\nabla} Q[X_\theta]||^2) d\theta \ , \quad Q[X] = \inf\ _{\sigma \in \mathcal{S}_N}\ ||X - A_\sigma||^2/2,$$

from which we may recover, through the least action principle, a dissipative version of the discrete VMA system (35,36)

$$\frac{d^2 X_\theta}{d\theta^2} = X_\theta - A_{\sigma_{opt}} \ , \quad \sigma_{opt} = \text{Arginf}\ _{\sigma \in \mathcal{S}_N}\ ||X_\theta - A_\sigma||^2$$

which, in particular, includes sticky collisions in the case $d = 1$.

(*) According to L. Ambrosio (private communication).

13 Appendix: Approximation of a generalized flow by introduction of an extra dimension

We consider here an optimal solution of the generalized geodesic problem, (c, m), where the domain is the unit square $D = [0,1]^2$, for times between $t_0 = 0$ and $t_1 = T$. We want to find a family of smooth divergence-free field $w_t^n(x) \in \mathbb{R}3$ compactly supported, and the associated volume-preserving flow $g_t^n(x)$ defined by $\partial g_t^n(x) = w_t^n(g_t^n(x))$, $g_0^n(x) = x$, such that

$$\int_{[0,T]\times\mathcal{A}} f(t, g_t^n(X_0(a), a)\mu(da)dt \to \int_{t,x,a} f(t,x,a)c(t,x,a), \quad \forall f \in C^0(Q), \tag{42}$$

$$\int_{[0,T]\times D} |w^n(t,x)|^2 dt dx \to \int_{t,x,a} |v(t,x,a)|^2 c(t,x,a) \tag{43}$$

in the interior of $[0,T] \times D$ family of maps in $X(t,\cdot) \in OPERASDiff_+(D \times [0,\varepsilon])$ which in some sense is close to (c, m) and which has almost the same energy:

$$\int_0^1 \int_{D\times[0,\varepsilon]} |\partial_t X(t,\tilde{a})|^2 d\tilde{a} dt \sim K(c,m).$$
$$\delta(\tilde{x} - X(t,\tilde{a})) \sim c(t,x,a)$$

Here $\sim$ means that we can find a left hand side which is arbitrarily close to the right hand side with respect to the “usual” norm.

Remark 54. • The above approximation is relevant, for example when we want to model the movements of the atmosphere, or of the oceans, where one dimension (the “vertical” one) is negligible with respect to the others. It can be interpreted as saying that, when dealing with fluid movements in 2-dimensional domains which arise as limits of 3-dimensional domains for one dimension going to zero, the natural setting is the one of generalized flows.

- The approximation theorem which we describe here is somehow unsatisfactory from the point of view of approximating generalized flows with classical flows, in the same sense in which the Nash embedding theorem with codimension 2 is unsatisfactory compared with the theorem in codimension 1: in our case we would like to do a similar, physically relevant, approximation, without introducing extra dimensions. Some approximation results which avoid extra dimensions are already present in [REF Shnirelman, Ambrosio-Figalli] but they are not fully satisfactory, [EXPLAIN].

To prove the above approximation, we proceed in several steps.

Step 1: mollification

We first prove the following approximation result:

Proposition 55. *Let $Q = [0,T] \times D \times \mathcal{A}$ where the space of labels is chosen to be $\mathcal{A} = \mathbb{T} = \mathbb{R}/Z$ with the Lebesgue measure as μ, and D is the $d-1$ dimensional cube $D = [0,1]^{d-1}$. Let $(c, m = cv) \in E' = \mathcal{C}(Q\,; \mathbb{R} \times \mathbb{R}d - 1)$ be such that*

$$\int_a c = 1, \quad \partial_t c + \nabla \cdot (cv) = 0, \;\; cv \parallel \partial D, \;\; K(c,m) = \frac{1}{2} \int |v|^2 \, c < +\infty.$$

Then, we can find a sequence $(c_n, m_n = c_n v_n)$, made of smooth functions on Q, with values in $\mathbb{R} \times \mathbb{R}d$, such that the following hold:

- *$(c_n, m_n) \rightharpoonup (c, m)$, for the the weak-$*$ convergence of measures (up to the boundary of Q);*
- *$c_n \geq \frac{1}{n}$ and $\int_\alpha c_n d\alpha = 1$;*
- *$\partial_t c_n + \nabla_x \cdot \; m_n = 0$ in $]0,1[\times \mathring{D} \times [0,1]$;*
- *$K(c_n, m_n) \leq K(c,m) + o(1)$ as $n \to +\infty$.*

Proof. The proof will consist in extending the variables t, x, a to $\mathbb{R} \times \mathbb{R}d - 1 \times \mathbb{T}$, then shrink a little bit the domain in the t, x-variables, and finally perform a mollification. All such steps will keep the action arbitrarily close to $K(c,m)$.

Extending the variables. We extend (c,m) outside $[0,T] \times D \times \mathbb{T}$ so that:

- for $x \notin D$ there holds $c = 1$, $m = 0$: therefore automatically (CE) holds, since $m \parallel \partial D$
- for $x \in D$ and $t < 0$ the particles stay frozen, i.e. we define $c(t,x,a) = \delta(x - X_0(a))$, $v(t,x,a) = 0$,
- for $x \in D$ and $t > T$ the particles also stay frozen, this time in their final position: we define $c(t,x,a) = \delta(x - X_T(a))$, $v(t,x,a) = 0$.

Retracting the space-time domain to a smaller one. We now shrink the variables t, x so that $[0, T] \times D = [0, 1]^{d-1}$ shrinks to $[\varepsilon, (1-\varepsilon)T] \times [\varepsilon, 1-\varepsilon]^{d-1}$ and we suitably change (c, m). This does not change neither the continuity equation nor the incompressibility condition and only affects $K(c, m)$ by a factor $1 + o(1)$, as $\varepsilon \to 0$). The improvement is that now $m = 0$ vanishes ouside of $[\varepsilon, (1-\varepsilon)T] \times [\varepsilon, 1-\varepsilon]^{d-1} \times \mathbb{T}$.

Positivity of c_ε and convolution. We first perform a convex interpolation by substituting for (c, m) the new pair $(\varepsilon + (1-\varepsilon)c, (1-\varepsilon)m)$, which maxes $c \geq \varepsilon > 0$, without affecting the continuity equation and the incompressibility condition and $K(c, m)$ is again modified by a factor $1 + o(1)$.

Next,we use a symmetric and positive kernel to mollify (c, m) in each variable separately. This does not modify the incompressibility condition and the continuity equation and we keep $m = 0$ outside of $[\varepsilon, (1-\varepsilon)T] \times [\varepsilon, 1-\varepsilon]^{d-1} \times T$ (after dividing ε by 2). So doing, we have reduced $K(c, m)$, since K is convex. Let us emphasize that, at each step, we have only performed small, controlable, modifications of (c, m) in the weak-* sense of measures. $\square$

Step 2: construction of a suitable velocity field with an extra dimension

Now we take $(c_n, m_n = c_n v_n)$, for some n big enough, as in the previous subsection, and we denote it by $(c, m = cv)$. We now consider the new spatial domain $D_d = D_{d-1} \times \mathbb{T} = [0, 1]^d$ whose variable will be denoted by $(x, z) \in D \times \mathbb{T}$. The new vertical coordinate $z \in \mathbb{T}$ is going to substitute, in a non-trivial way, for the label variable $a \in \mathbb{T}$.
To pass from the label $a \in \mathbb{T}$ to the vertical variable $z \in \mathbb{T}$ representing the "extra dimension", we consider the monotone rearrangement map $R(t, x, \cdot) : [0, 1] \to [0, 1]$ sending $c(t, x, \cdot)$ to the 1D Lebesgue measure on $[0, 1]$. More precisely, we implicitly define the function $z \mapsto R(t, x, z)$,

$$\int_0^{R(t,x,z)} c(t, x, a)da = z, \quad \forall z \in [0, 1], \tag{44}$$

which implies that, for all bounded Borel function f, and for all $(t, x) \in [0, T] \times D_{d-1}$,

$$\int_0^1 f(R(t, x, z))dz = \int_0^1 f(a)c(t, x, a)da.$$

Since c is smooth and bounded away from 0, $R(t, x, z) - z$ is a smooth function, periodic in z, with $R(t, x, 0) = 0$, $R(t, x, 1) = 1$ and $\partial_z R > 0$. We then define a smooth divergence-free vector field depending on time

$$(t, x, z) \in [0, T] \times D_{d-1} \times \mathbb{T} \to (u(t, x, z), w(t, x, z)) \in \mathbb{R}d \times \mathbb{R}$$

$$
\begin{aligned}
u(t,x,z) &= v(t,x,R(t,x,z)) \text{ (where } v = \frac{m}{c} \text{)} \\
w(t,x,z) &= -\int_0^z (\nabla_x \cdot u)(t,x,\tilde{z})d\tilde{z},
\end{aligned}
$$

The fact that $\partial_z w + \nabla_x \cdot u = 0$ is clear from the definition of w, and means exactly that (u, w) is divergence-free.

We introduce the volume-preserving flow (ξ_t, η_t) generated on $D_{d-1} \times \mathbb{T}$ by (u, w) through: Notice that, since $w = 0$ at $z = 0$ and $z - 1$ By definition of R, we have for all continuous function f

$$
\begin{aligned}
&\partial_t \xi = u(t,\xi,\eta) \\
&\partial_t \eta = w(t,\xi,\eta) \\
&(\xi,\eta)(0,x,z) = (x,z)
\end{aligned}
$$

on $\mathbb{R} \times \mathbb{R}d - 1 \times \mathbb{T}$

$$
\int_{t,x,a} f(t,x,a)c(t,x,a) = \int_{t,x,z} f(t,x,R(t,x,z)).
$$

Since (ξ_t, η_t) is a volume-preserving diffeormorphism. this can be also written

$$
\int_{t,x,a} f(t,x,a)c(t,x,a) = \int_{t,x,z} f(t,\xi_t(x,z),\tilde{R}(t,x,z))
$$

where

$$
\tilde{R}(t,x,z) = R(t,\xi_t(x,z),\eta_t(x,z)).
$$

Similarly, by definition of R and u,

$$
\int_{t,x,a} f(t,x,a)m(t,x,a) = \int_{t,x,z} f(t,x,R(t,x,z))v(t,x,R(t,x,z))
$$

$$
= \int_{t,x,z} f(t,x,R(t,x,z))u(t,x,z) = \int_{t,x,z} f(t,\xi_t(x,z),\tilde{R}(t,x,z))u(t,\xi_t(x,z),\eta_t(x,z))
$$

$$
= \int_{t,x,z} f(t,\xi_t(x,z),\tilde{R}(t,x,z))\frac{d}{dt}\xi_t(x,z).
$$

Now, let us use that $(c, m) = (c_n, m_n)$ almost satisfies condition (CE), as $n \to \infty$, namely, for all sufficiently smooth function $f(t, x, a)$,

$$
\int_{t,x,a} \partial_t f c + \nabla_x f \cdot m = BT_n(f) \sim BT(f), \quad n \to \infty,
$$

where

$$
BT_n(f) = \int_{x,a} f(T,x,a)c(T,x,a) - f(0,x,a)c(0,x,a),
$$

$$
BT(f) = \int_a f(T,X_T(a),a) - f(0,X_0(a),a).
$$

Using the new expression of (c, m) we have just obtained, we get

$$BT_n(f) = \int_{t,x,z} (\partial_t f)(t, \xi_t(x,z), \tilde{R}(t,x,z)) + (\nabla_x f)(t, \xi_t(x,z), \tilde{R}(t,x,z)) \frac{d}{dt}\xi_t(x,z)$$

$$= \int_{t,x,z} \frac{d}{dt}[f(t, \xi_t(x,z), \tilde{R}(t,x,z))] - (\partial_a f)(t, \xi_t(x,z), \tilde{R}(t,x,z))\partial_t \tilde{R}(t,x,z)$$

$$= \int_{x,z} [f(T, \xi_T(x,z), \tilde{R}(T,x,z)) - f(0, x, \tilde{R}(0,x,z))]$$

$$- \int_{t,x,z} (\partial_a f)(t, \xi_t(x,z), \tilde{R}(t,x,z))\partial_t \tilde{R}(t,x,z).$$

Let us first use test functions f that vanish at $t = 0$ and $t = T$. We then get

$$0 = \int_{t,x,z} (\partial_a f)(t, \xi_t(x,z), \tilde{R}(t,x,z))\partial_t \tilde{R}(t,x,z),$$

The right-hand side can be written, using the definition of $\tilde{R}$

$$\int_{t,x,z} (\partial_a f)(t, \xi_t(x,z), R(t, \xi_t(x,z), \eta_t(x,z)))(D_t R)(t, \xi_t(x,z), \eta_t(x,z)),$$

(where $D_t R$ is a short notation for $(\partial_t + u \cdot \nabla_x + w\partial_z)R$) which is nothing but

$$\int_{t,x,z} (\partial_a f)(t, x, R(t,x,z)) D_t R(t,x,z)$$

(since (ξ_t, η_t) is a volume-preserving diffeoorphism). By setting $g(t,x,z) = f(t, x, R(t,x,z)$, so that

$$\partial_z g(t,x,z) = (\partial_a f)(t, x, R(t,x,z))\partial_z R(t,x,z),$$

we deduce

$$\int_{t,x,z} \partial_z g(t,x,z) \frac{D_t R(t,x,z)}{\partial_z R(t,x,z)},$$

which is possible only if $D_t R(t,x,z) = \partial_z R(t,x,z)\beta(t,x)$ for some function $\beta(t,x)$. In other words

$$(\partial_t + u \cdot \nabla_x + (w - \beta)\partial_z)R = 0.$$

Since $w(t,x,z)$ has been defined only up to a function $\hat{w}(t,x)$, we may adjust $\hat{w}(t,x)$ so that $\beta(t,x) = 0$ and get:

$$(\partial_t + u \cdot \nabla_x + w\partial_z)R = 0.$$

This means that $R(t, \xi_t(x,z), \eta_t(x,z)) = R(0,x,z)$ and widely simplifies the formulae we have obtained for (c, m). Indeed, we may now write

$$\int_{t,x,a} f(t,x,a)c(t,x,a) = \int_{t,x,z} f(t, \xi_t(x,z), R(0,x,z))$$

$$\int_{t,x,a} f(t,x,a)m(t,x,a) = \int_{t,x,z} f(t,\xi_t(x,z),R(0,x,z))\frac{d}{dt}\xi_t(x,z),$$

Going back to condition (CE), we have finally obtained for the time-boundary data

$$BT_n(f) = \int_{x,z} f(T,\xi_T(x,z),R(0,x,z)) - f(0,x,R(0,x,z))$$

$$\sim BT(f) = \int_a f(T,X_T(a),a) - f(0,X_0(a),a),$$

for all f, as $n \to \infty$.

Step 3: matching the time-boundary data

Let us now introduce smooth approximation for X_0 and X_T, respectively denoted by X_0^ε and X_T^ε, so that

$$\int_{[0,1]} |X_0^\varepsilon(a) - X_0(a)|^2 da \le \varepsilon^2, \quad \int_{[0,1]} |X_T^\varepsilon(a) - X_T(a)|^2 da \le \varepsilon^2 \ ,$$

we choose successively $f(t,x,a) = (T-t)|x - X_0^\varepsilon(a)|^2$ and $f(t,x,a) = t|x - X_T^\varepsilon(a)|^2$ and get

$$\lim_n \int_{x,z} |\xi_n(T,x,z) - X_T^\varepsilon(R_n(0,x,z))|^2 = \int_{[0,1]} |X_T^\varepsilon(a) - X_T(a)|^2 da \le \varepsilon^2 \ ,$$

$$\lim_n \int_{x,z} |x - X_0^\varepsilon(R_n(0,x,z))|^2 = \int_{[0,1]} |X_0^\varepsilon(a) - X_0(a)|^2 da \le \varepsilon^2 \ .$$

By the triangle inequality, we have

$$\sqrt{\int_{x,z} |x - X_0^\varepsilon(R_n(0,x,z))|^2} - \sqrt{\int_{x,z} |x - X_0(R_n(0,x,z))|^2}$$

$$\le \sqrt{\int_{x,z} |X_0^\varepsilon(R_n(0,x,z)) - X_0(R_n(0,x,z))|^2}$$

$$= \sqrt{\int_{[0,1]} |X_0^\varepsilon(a) - X_0(a)|^2 da} \ \le \ \varepsilon$$

(by construction of R_n). Similarly, we get

$$\sqrt{\int_{x,z} |\xi_n(T,x,z) - X_0^\varepsilon(R_n(0,x,z))|^2} \le \sqrt{\int_{x,z} |\xi_n(T,x,z) - X_0(R_n(0,x,z))|^2} + \varepsilon \ .$$

So, we can pass to the limit in ε and get

$$\int_{x,z} |\xi_n(T,x,z) - X_T(R_n(0,x,z))|^2 \to 0,$$

$$\int_{x,z} |x - X_0(R_n(0,x,z))|^2 \to 0 \ .$$

Next, we use the assumption there is, for each ε, a smooth function $h_\varepsilon : D \to D$ such that

$$\int_{[0,1]} |X_T(a) - h_\varepsilon(X_0(a))|^2 da \le \varepsilon^2.$$

Thus,

$$\sqrt{\int_{x,z} |\xi_n(T,x,z) - X_T(R_n(0,x,z))|^2} - \sqrt{\int_{x,z} |\xi_n(T,x,z) - h_\varepsilon(X_0(R_n(0,x,z)))|^2}$$

$$\le \sqrt{\int_{x,z} |X_T(R_n(0,x,z)) - h_\varepsilon(X_0(R_n(0,x,z)))|^2} = \sqrt{\int_a |X_T(a) - h_\varepsilon(X_0(a))|^2} \le \varepsilon.$$

Using that

$$\int_{x,z} |h_\varepsilon(x) - h_\varepsilon(X_0(R_n(0,x,z)))|^2 \le \mathrm{Lip}(h_\varepsilon)^2 \int_{x,z} |x - X_0(R_n(0,x,z)))|^2 \to 0,$$

we have obtained

$$\limsup_n \int_{x,z} |\xi_n(T,x,z) - h_\varepsilon(x)|^2 \le \varepsilon^2,$$

which can also be written

$$\limsup_n \int_{a,z} |\xi_n(T,X_0(a),z) - h_\varepsilon(X_0(a))|^2 \le \varepsilon^2,$$

By passing to the limit in ε, we have finally obtained:

Proposition 56.

$$\lim \int_{a,z} |\xi_n(T,X_0(a),z) - X_T(a)|^2 = 0. \tag{45}$$

Step 4: rescaling the vertical direction

We here simply show how to do the rescaling of the flow, changing the vertical space direction from an interval $[0,1]$ to a short interval $[0,\varepsilon]$: on $[0,\varepsilon]$ we define

$$\begin{cases} \tilde{u}(t,x,z) = u(t,x,z/\varepsilon) \\ \tilde{w}(t,x,z) = \varepsilon w(t,x,z/\varepsilon), \end{cases}$$

and from them we again obtain a flow $\tilde{\xi}, \tilde{\eta}$ as above. Then the action of such flow will be

$$\begin{aligned} 2K(\tilde{c}, \tilde{m}) &\sim \int |\partial_t \xi|^2 + \int |\partial_t \eta|^2 \\ &= \int |\tilde{u}|^2 + \int |\tilde{w}|^2 \\ &\sim \int |u|^2 + \varepsilon^2 \int |v|^2 \\ &\leq 2K(c, m) + o(1), \end{aligned}$$

while the previous estimates on the boundary conditions, as well as the continuity and incompressibility equations, continue to hold by easy computations.

14 Concrete examples

14.1 a 1-dimensional classical flow

We will give an example of a minimal geodesic in $OPERAVPM(D)$ between times $t_0 = 0$ and $t_1 = 1$, where the domain D is the 1-dimensional segment $[-1, 1]$.

The flow will be given by the function

$$X(t, a) = \begin{cases} t^{2/3} h(at^{-2/3}) \text{ if } a \in [-t^{2/3}, t^{2/3}] \\ a \text{ else} \end{cases}$$

where

$$h(a) = \begin{cases} 2\sqrt{a} - 1 \text{ if } 0 < a \leq 1 \\ 1 - 2\sqrt{-a} \text{ if } -1 \leq a < 0 \end{cases}$$

[FIGURE]

Then we claim that $X(t, \cdot)$ is minimizing the action $\int_{t,x} |\partial_t X|^2$ in the class $OPERAVPM(D)^{[0,1]} \cap \{X(t, \cdot) : X(0, \cdot) = id,\ X(1, \cdot) = h_1\}$. Moreover, we can show that if we denote as usual $c(t, x, a) = \delta(x - X(t, a)), m(t, x, a) = \partial_t X(t, a) c(t, x, a)$, then (c, m) is also an optimal generalized flow.

To show our claim we want to use the optimality criterion stating that if $X(t, \cdot) \in OPERAVPM(D)$ for all times t and if for almost all $a \in D$, the path $X(\cdot, a)$ minimizes the integral $\int_t [|\dot{\xi}(t)|^2 - p(t, \xi(t))] dt$ in the class $\{\xi \text{ smooth and } \xi(0) = a, \xi(1) = h_1(a)\}$ where p is the pressure field of X, then X is a minimizing flow. We state the general result which we want to use, as a proposition:

Proposition 57. *Consider an incompressible flow $X(t,a,\omega)$ on a domain D, where $\omega \in [0,1]$ is a parameter, whose pressure field is $p(t,x)$. In other words, calling*

$$\begin{aligned} c(t,x,a) &= \int \delta(x - X(t,a,\omega))d\omega \\ m(t,x,a) &= \int \partial_t X(t,a,\omega)\delta(x - X(t,a,\omega))d\omega, \end{aligned}$$

We want the following conditions to hold:

(I) $\int f(X(t,a,\omega))da\, d\omega = \int f(x)dx$ *for all test functions* f

(II) $\partial_{tt}X(t,a,\omega) = -\nabla_x p(t, X(t,a,\omega))$

(III) $t \mapsto X(t,a,\omega)$ *minimizes the functional*

$$I(\xi) = \int_{t_0}^{t_1} \left[\frac{|\dot{\xi}(t)|^2}{2} - p(t,\xi(t)) \right] dt,$$

among all $\xi : [t_0, t_1] \to D$ *satisfying* $\xi(t_0) = X(t_0,a,\omega)$, $\xi(t_1) = X(t_1,a,\omega)$.

Then the flow (c,m) has minimal action in the class of the incompressible generalized flows connecting $c(t_0,x,a)$ and $c(t_1,x,a)$ on a time interval of length $t_1 - t_0$.

We may consider our flow $X(t,a)$ as a flow of the form $X(t,a,\omega)$ which does not depend on the parameter ω. Therefore, in order to apply the above proposition, we just have to perform the following three computations:

- *We show that $X_t = X(t,\cdot)$ is volume-preserving, i.e., condition (I) of Proposition 57 holds.* To do this, we take a test function f and we compute:

$$\begin{aligned} \int_{-1}^{1} f(X_t(a))da &= \int_{[-1,1]\setminus[-t^{2/3},t^{2/3}]} f(a)da + \int_{-t^{2/3}}^{t^{2/3}} f(t^{2/3}h(at^{-2/3}))da \\ &= \int_{[-1,1]\setminus[-t^{2/3},t^{2/3}]} f(a)da + t^{2/3}\left[\int_0^1 f(t^{2/3}(2\sqrt{b}-1))db + \int_{-1}^{0} f(t^{2/3}(1-2\sqrt{-b}))db \right] \\ &= \int_{[-1,1]\setminus[-t^{2/3},t^{2/3}]} f(a)da + \frac{1}{2}\int_{-t^{2/3}}^{t^{2/3}} f(c)(t^{-2/3}c+1)dc + \frac{1}{2}\int_{-t^{2/3}}^{t^{2/3}} f(c)(1-t^{-2/3}c)dc \\ &= \int_{-1}^{1} f(a)da, \end{aligned}$$

  as wanted.

- *We compute the pressure field, defnined by condition (II) of Proposition 57.* We have

$$\partial_{tt}X(t,a) = -\frac{2}{9}t^{-4/3} \cdot \begin{cases} 2t^{-1/3}\sqrt{a} - 1 & \text{if } a \in]0, t^{2/3}] \\ 1 - 2t^{-1/3}\sqrt{-a} & \text{if } a \in [-t^{2/3}, 0[\\ 0 & \text{else} \end{cases}$$

We would like to say that $\partial_x p(t,x) = -\partial_{tt} X_t \circ X_t^{-1}(x)$, but actually X_t is not injective:

$$X_t^{-1}(\{x\}) = \begin{cases} \left\{ \pm t^{-2/3} \left(\frac{x \pm t^{2/3}}{2} \right)^2 \right\} & \text{for } x \in [-t^{2/3}, t^{2/3}] \\ \{x\} \text{ else} \end{cases}$$

We have however that the spatial derivative of p is well defined, and more precisely:

$$\{\partial_x p(t,x)\} = (-\partial_{tt} X_t)(X_t^{-1}(\{x\})) = \begin{cases} \left\{ \frac{2}{9} t^{-2} a \right\} & \text{if } x \in [-t^{-2/3}, t^{2/3}] \\ \{0\} & \text{else} \end{cases}$$

Integrating the above from $p(\cdot, -1) := 0$ (which we are allowed to fix, being p defined only up to the summation of a t-dependent function), we obtain

$$p(t,x) = -\frac{1}{9t^2}(t^{4/3} - x^2)_+. \tag{46}$$

- *We show the optimality condition (III) of Proposition 57.* The functional of Proposition 57 in our case reads:

$$\int_0^1 \left[\frac{|\dot{\xi}(t)|^2}{2} - \frac{1}{9t^2}(t^{4/3} - \xi^2(t))_+ \right] dt$$

The trajectory $t \mapsto (X(t,a), t)$ in $[0,1] \times ([-1,1] \setminus \{0\})$ is a vertical segment until it reaches the graph of the convex function $t(a) = a^{3/2}$, and then stays always inside the epigraph of this function. $X_a(t) = X(\cdot, a)$ is a piecewise smooth solution of the equation $\partial_{tt} X_a(t) = -\partial_x p(t, X_a(t))$, and in particular it is a critical point of the action. We also observe that $X(\cdot, a)$ is a C^1-regular function, and in particular the derivative at $t = |a|^{3/2}$ is zero on both sides of this discontinuity point. Asking that our functional is increasing along all perturbations of ξ by functions ζ which vanish at times $0, 1$ is then expressed by asking that the following holds:

$$I := \int_0^1 \left[\frac{1}{2} \left| \frac{d(\xi + \zeta)(t)}{dt} \right|^2 - p(t, \xi(t) + \zeta(t)) \right] dt \geq \int_0^1 \left[\frac{1}{2} \left| \frac{d\xi(t)}{dt} \right|^2 - p(t, \xi(t)) \right] dt := I_0 \tag{47}$$

In order to do the above estimate, we observe that the pressure satisfies

$$\partial_{xx} p(t,x) \leq \frac{2}{9t^2},$$

so we have the semiconcavity inequality

$$p(t, \xi(t) + \zeta(t)) \leq p(t, \xi(t)) + \partial_x p(t, \xi(t)) \zeta(t) + \frac{\zeta^2(t)}{9t^2}.$$

Inserting this in the right hand side of (47) we obtain

$$I \leq I_0 + I_1 + I_2,$$

where (since

$$
\begin{aligned}
I_1 &= \int_0^1 \left[\frac{d\xi(t)}{dt}\frac{d\zeta(t)}{dt} - \partial_x p(t,\xi(t))\zeta(t)\right] dt \\
I_2 &= \int_0^1 \left[\frac{1}{2}\left|\frac{d\zeta(t)}{dt}\right|^2 - \frac{2\zeta(t)}{9t^2})\right] dt.
\end{aligned}
$$

and the fact that $I_2 \geq 0$ is given by the Hardy inequality [1] :

$$
\int_0^\infty \left[\left|\frac{d\zeta(t)}{dt}\right|^2 - \frac{\zeta^2(t)}{4t^2}\right] dt,
$$

while by using the fact that $\partial_{tt}X(t,x) = -\partial_x p(t, X(t,x))$ and a change of variables we see that also $I_1 = 0$ for the trajectories of our flow. We observe that

[FIGURE:trajectories of a point a]

Remark 58. The pressure gradient $\partial_x p(t,x)$ has a discontinuity at the interface $\{t = |a|^{3/2}\}$, while our flow is still optimal. This shows us that the most we can achieve in terms of the regularity of pressure, without assuming particular properties on the boundary data, is semiconcavity.

Remark 59. The flow $X(t,\cdot)$ constructed above is singular at the origin at time $t = 0$, and for $a \in [-1,1]\setminus\{0\}$ the pressure $p(0,a)$ is zero, and actually as $t \to 0$, the pressure field tends to a Dirac mass: its integral is indeed constantly equal to $4/27$ since

$$
\int_{-1}^1 [(t^{4/3} - x^2)_+/9t^2]dx = 1/9\int_{-1}^1 (1-\xi^2)d\xi = 4/27,
$$

while its support concentrates to the origin.
Then the pressure will become nonzero at point $a \neq 0$ only at time $t = |a|^{3/2}$, so we have what is called a *finite speed of propagation* of the motion in the fluid, which may seem unexpected for an incompressible fluid.

Remark 60. By its self-similarity, the flow is minimal on all positive time intervals, which means that we may construct an *eternal solution* by extending the flow by self-similarity.

[1] Usually Hardy's inequality is stated in the form

$$
\int_0^\infty \left(\tfrac{1}{t}\int_0^t f(\tau)d\tau\right)^p dt \leq (\tfrac{p}{p-1})^p \int_0^\infty f^p(t)dt,
$$

which holds as soon as $p > 1$ and $f \geq 0$. Here we just put $\zeta(t) := \int_0^t f(\tau)d\tau$ and we take $p = 2$. The resulting condition on ζ will be that $\zeta(0) = 0$ and that ζ is increasing.

Remark 61. This particular flow is the only self-similar flow with a singularity at the origin at time zero. Indeed, suppose we have a flow of the form

$$X(t,a) = X_{\beta,\gamma}(t,a) = \begin{cases} a & \text{if } |a| > t^{-\beta} \\ \bar{X}(t^\beta a)t^\gamma & \text{if } |a| < t^{-\beta} \end{cases}$$

Then the change of variable formula (where we suppose that $\bar{X}(t,\cdot)$ is incompressible) gives for any test function φ

$$\int \varphi(X(t,a))da = \int_{|a|>t^{-\beta}} \varphi(a)da + \int_{|b|<1} \varphi(\bar{X}(b)t^\gamma)t^{-\beta}db = \int_{|a|>t^{-\beta}} \varphi(a)da + \int_{|b|<1} \varphi(bt^\gamma)t^{-\beta}db,$$

whence we obtain $\gamma = -\beta$. On the other hand, in order to impose the conservation of kinetic energy we can compute (on the zone $|a| < t^{-\beta}$ where it is nonzero) the velocity field

$$\partial_t X(t,a) = \beta t^{-\beta-1}\left[at^\beta \bar{X}'(t^\beta a) - \bar{X}(t^\beta a)\right],$$

so the kinetic energy is

$$\int_{|a|<t^{-\beta}} |\partial_t X(t,a)|^2 da = t^{-3\beta-2}\int_{|b|<1} |b\bar{X}'(b) - X(b)|^2 db,$$

and in order for this quantity to be constant in time, we need $\beta = -2/3$. The only remaining freedom for us now is the choice of $\bar{X}$ [COMPLETE PROOF]

14.2 Lifting to an extra dimension

We may now lift our flow to an extra dimension by introducing a new variable $z \in [0,1]$ and a flow in the two dimensions x, z can then be defined once we define a divergence-free vector field $(u(t,x,z), w(t,x,z)) \in \mathbb{R}2$. Since we need the incompressibility condition $\partial_x u + \partial_z w = 0$, we may suppose without loss of generality that

$$\begin{cases} u = \partial_z \psi(t,x,z) \\ w = -\partial_x \psi(t,x,z) \end{cases}$$

Such kind of function ψ is called *stream function*. We start by defining

$$\psi(1,x,z) = \begin{cases} \frac{z(x-1)}{3} & \text{if} \quad 0 < z < \frac{1+x}{2} \\ \frac{(z-1)(x+1)}{3} & \text{if} \quad \frac{1+x}{2} \le z < 1 \end{cases}$$

and then we extend by letting $\psi(t,x,z) = t^{-1/3}\psi(1, t^{-2/3}x, z)$ for $|x| < t^{2/3}$ and $\psi(t,x,z) = 0$ for $|x| \in [t^{2/3}, 1]$. The levelsets of $\psi(t,\cdot,\cdot)$ are easily seen to be pieces of hyperbolas on each one of the sets

$$\begin{gathered} \{(x,z) : |x| < t^{2/3}, 0 < z < \tfrac{1+t^{-2/3}x}{2}\} \\ \{(x,z) : |x| < t^{2/3}, \tfrac{1+t^{-2/3}x}{2} < z < 1\}, \end{gathered}$$

while for $|x| > t^{2/3}$, ψ is constant. In other words, we have again that the particle starting at (a, z) stays frozen until time $t = a^{3/2}$ and then starts to move, first sliding along the boundary of the rectangle $[(|a|, 0), (-|a|, 0), (-|a|, 1), (|a|, 1)]$ at time $t = a^{3/2}$, and then taking a more curved orbit in a clockwise direction, as time passes. [FIGURE: levelsets of ψ]
It is clear that on the other side the velocity field (u, w) is tangent to the levelsets of ψ, and if we call $\xi = t^{-2/3}x$ the rescaled horizontal coordinate then the stream function is continuous at the interface s $\xi = 1, \xi = -1, z(1 + \xi)/2$, and one obtains the velocity field

$$(u, w) = \begin{cases} \left(\frac{t^{-1/3}(\xi+1)}{3}, -\frac{t^{-1}z}{3}\right) & \text{if } 0 < z < \frac{1+\xi}{2} \text{ and } |\xi| \leq 1 \\ \left(\frac{t^{-1/3}(\xi-1)}{3}, -\frac{t^{-1}(z-1)}{3}\right) & \text{if } \frac{1+\xi}{2} < z < 1 \text{ and } |\xi| \leq 1 \\ (0, 0) & \text{if } |\xi| > 1. \end{cases}$$

We now observe that considering just the x-projections of our particles brings us to the previous flow (in other words, *we have lifted our previous generalized flow to a* 2*-dimensional incompressible flow*). To show this, we first observe that u (i.e. the horizontal component of the velocity) does not depend on the vertical variable z; we have to verify that

$$u(t, x) = \left[\partial_t X_t \circ X_t^{-1}\right](x), \text{ where we put } X_t(a) = X(t, a). \tag{48}$$

Since we already have calculated the expression of $X_t^{-1}(x)$, the above formula follows easily.

The Euler equations in this case can be verified to hold in the following form:

$$\begin{cases} \partial_x u + \partial_z w = 0 \\ \partial_t u + \partial_x(u^2) + \partial_z(uw) = -\partial_x p(t, x) \\ \partial_t w + \partial_x(uw) + \partial_z(w^2) = -\partial_z p = 0 \end{cases}$$

and this kind of system where the pressure is constant a given direction (in this case it is the vertical one) is called a *hydrostatic limit*.

After defining such flow, we can vertically rescale the domain to $[-1, 1] \times [0, \varepsilon]$ by the vertical change of variable $z, w \mapsto \varepsilon z, \varepsilon w$. the first two of the Euler ewuations then remain formally unaltered, while the third one changes into

$$\varepsilon(\partial_t w + \partial_x(uw) + \partial_z(w^2)) = -\partial_z p.$$

which gives a smooth solution approximating the one in the previous example on $[-1, 1]$, since it means that the pressure becomes constant in the z direction, as $\varepsilon \to 0$.

Remark 62. A reasonable conjecture is that any generalized solution on a domain D can be viewed as a hydrostatic limit of a "lift" of it to a solution on the same domain with an extra dimension, $D \times [0, 1]$. The difficulty in this conjecture is that one cannot a priory assume sufficient regularity in order to define the lift in a differential geometric manner.

15 Appendix: Measure preserving maps and density theorems

In this appendix, measure preserving maps are studied in a relatively general framework. Let us start with

Definition 63. *Let X and Y be two topological spaces. Let α and β be two Borel probability measures respectively defined on X and Y. We say that a map $\phi : X \to Y$ transports (X, α) onto (Y, β) or that β is the image of α by ϕ if, for all borel subset B of Y, $\phi^{-1}(B)$ is a Borel set in X and $\alpha(\phi^{-1}(B)) = \beta(B)$. When $X = Y$ and $\alpha = \beta$, we say that ϕ is a measure preserving map (MPM).*

Remarks and examples

1) An equivalent definition is given by : for all Borel function f which is $\beta-$ integrable on Y, $x \to f(\phi(x))$ is Borel and $\alpha-$ integrable on X and

$$\int_X f(\phi(x))d\alpha(x) = \int_Y f(y)d\beta(y). \tag{49}$$

2) Of course, the definition can be extended to abstract measure spaces.
3) In the case $X = Y = [0, 1]$, $\alpha = \beta = |\cdot|$, where $|\cdot|$ denotes the Lebesgue measure, some examples of measure preserving maps are given by

$$\phi(x) = x \tag{50}$$

$$\phi(x) = x + \frac{1}{2} \mod.1 \tag{51}$$

(which is discontinuous),

$$\phi(x) = 1 - x \tag{52}$$

(which is orientation reversing),

$$\phi(x) = \min(2x, 2 - 2x) \tag{53}$$

(which is not one-to-one).

4) A remarkable theorem (see [**?**] for example) asserts that if X is a separable complete metric space and no point in X has α positive measure ($\alpha(x) = 0, \forall x \in X$), then there is a map $\phi : X \to Y = [0, 1]$ that transports α to the Lebesgue measure on $[0, 1]$. (The idea of the construction is quite simple. Let (a_n), $n = 1, 2, ...$ be a dense sequence in X. Let rescale the distance d on X so that the diameter of X is one. To each point x in X, we associate the sequence $d(x) = (d(x, a_n)) \in [0, 1]^{\mathbf{N}}$, which provides a kind of system of coordinates in

X. Then, we use binary coding to write $d(x)$ as a point in $(\{0,1\}^{\mathbf{N}})^{\mathbf{N}}$, which is in one-to-one correspondance with $(\{0,1\}^{\mathbf{N}})$ and leads us back to $[0,1]$, through binary decoding. This establishes a correspondance ϕ between X and $[0,1]$. Further refinements are needed to make it one-to-one (in the almost everywhere sense). Then, it is easy to modify ϕ, by composition (using the property that no point in X has positive μ measure), to enforce (49). Of course, such a construction deserves to be done very carefully.) So, in some sense, from the measure theoretic point of view, there is no essential difference between (X,α) and $([0,1],|\cdot|)$ as long as X is metric, separable, complete without any point of positive α measure.

Remark

Of course, whenever X, Y are two finite sets of same cardinal N, both equipped with the counting measure, then the set of all measure preserving maps is isomorphic to the symmetric group $\mathfrak{S}_N$.

15.1 Smooth measure preserving maps

Subsequently, we consider the case when $X = Y = D$ is the closure of a bounded open set with Lipschitz boundary in $\mathbf{R}^d$ and $\alpha = \beta$ is the $d-$ dimensional Lebesgue measure, denoted by $|\cdot|$ and normalized so that $|D| = 1$. Typically D is the unit hypercube $[0,1]^d$. The set of all measure preserving maps (MPM) is denoted by $S = S(D)$, where S stands for *semi*−group. Indeed, by definition, S equipped with the usual composition rule is a semi-group, but not a group (due to the presence of obviously non invertible elements, such as example (53) in the simplest case $D = [0,1]$). This set S can be seen as a subset of the Lebesgue space $L^p(D,\mathbf{R}^d)$ of all p-integrable maps from D into $\mathbf{R}^d$ (the ambiant space), and more specifically as a closed subset of a sphere, whatever is the value of $p \in [1,+\infty]$. (Let us just recall that every Lebesgue measurable function is almost everywhere equal to a Borel function, so that there is no problem to define S as a subset of L^p.)

Remark

It is clear that S is closed and contained in a sphere.

Let us now consider more restrictive definitions of measure preserving maps, requiring some smoothness.

Let us first consider the vector space V of all time dependent C^∞ vector fields on D,

$$(t,x) \in [0,1] \times D \to v(t,x) \in \mathbf{R}^d,$$

compactly supported in the interior of $[0,1] \times D$ and divergence-free :

$$\nabla .v(t,x) = \sum_{i=1}^{d} \partial_{x_i} v_i(t,x) = 0.$$

Let us denote by $g_t(v)(x)$ the solution at time t of the ODE $dx/dt = v(t,x)$ with x as initial condition at $t = 0$. Because v is smooth and compactly supported, for all t, $g_t(v)$ is a C^∞ orientation preserving diffeomorphism of D, leaving a neighbourhood of the boundary ∂D pointwise unchanged. Since v is divergence free, the jacobian determinant of $g_t(v)$ is identically equal to one, because of the general identity

$$\partial_t \log \det(\partial_x g_t(v)(x)) = (\nabla .v)(t, g_t(v)(x)), \tag{54}$$

valid for all smooth vector fields v. In particular, $g_t(v)$ is (Lebesgue) measure preserving, because of the change of variable formula

$$\int_D f(\phi(x)) \det(\partial_x g_t(v)(x)) dx = \int_D f(x) dx, \ \ \forall f.$$

Thus,

$$G_0 = \{g_1(v) \quad v \in V\}.$$

defines a subset of the group G of all diffeomorphisms with unit jacobian determinant. Actually, G_0 is a subgroup of G. Clearly, G_0 is also a subset of the semigroup S of all measure preserving maps, which already contains G.

Exercise

Show that indeed G_0 is a group. Describe G_0 in the case $D = [0,1]$, $d = 1$.

Exercise

Prove (54) using that $\det(I + A) = 1 + \mathrm{tr(A)} + \mathrm{O(A^2)}$.

Exercice

Let ϕ be a Lipschitz map $D \to D$ belonging to S. Show that ϕ must satisfy

$$\sum_{x;\ \phi(x)=y} |\det \partial_x \phi(x)|^{-1} = 1,$$

for almost every $y \in D$. (To do a detailled and precise proof, the use of [129] is recommanded.) Can such map be smooth (at least C^1) without being one-to-one?

15.2 Density of smooth measure preserving maps

Clearly, S is a much larger set of maps than G_0. (The case $D = [0,1]$ is a striking example, since, then, G_0 is reduced to the identity map.) However, as shown in this section, from the point of view of L^p topologies, for $p < +\infty$, G_0 is dense in S as soon as $d \geq 2$. To make the proof as simple as possible we assume D to be the unit hypercube.

Theorem 2. *Let $D = [0,1]^d$ and $d \geq 2$, then S is the closure of G_0 in the space $L^p(D, \mathbf{R}^d)$, for all $p \in [1, +\infty[$.*

Remarks

1) Clearly $d \geq 2$ is needed !
2) For finer topologies than the L^p one ($p < +\infty$), S is usually strictly larger than the closure of G_0. This is obvious for the C^1 topology which preserves the unit jacobian determinant pointwise. Sobolev topologies, which occurs naturally in the theory of incompressible elasticity, such as $W^{1,p}$, preserve the unit jacobian determinant in the almost everywhere sense, at least for p large enough. In the very special case $d = 2$, the C^0 topology is almost sufficient to preserve the unit jacobian determinant. (This is in fact related to symplectic topology.) There has been a lot of researches related to these questions (let us quote a few names among others, at least in the field of Calculus of Variations, such as J. Ball, F. Dacorogna, S. Müller, T. Sverak, L. Tartar, and some related work by Coifman-Lions-Meyer, F. Hélein, C. Viterbo etc... as well as the book by Arnold and Khesin [12]...) The L^2 topology is too weak to preserve the unit jacobian determinant. As a matter of fact, orientation reversing maps such as $(x_1, x_2) \to (x_1, 1 - x_2)$ on the unit square can be approximated by elements of G_0 in L^p norm for $p < +\infty$, as we shall see.

15.3 Proof of the density theorem

There are several possible proofs of this "folklore" density result. The following one (due to the author but unpublished), does not differ much from the one provided in Neretin's paper [198].

15.3.1 Measure preserving maps and permutations

Usually, density results are proved using regularization techniques such as convolution. Here S is not a vector space and convolution cannot be used straightforwardly. Of course, since G_0 is formally a Lie group with a Lie algebra made up of smooth divergence free vector fields compactly supported in the interior of D, a natural idea would be to look for a vector space of generalized divergence free fields, to which convolution could be applied, that would generate S by integration. But there is no obvious space of that type (although the theory discussed in the second part of the text solves this problem in some sense). So, we are going to follow a completely different track relying on the approximation of S by a discrete group, the group of permutations. Indeed, at the discrete level, as D is a finite set of m elements with the counting measure, S can be identified to the group of the permutations of the m first integers. So, to approximate S, it is natural to introduce, for each integer $n \geq 0$, the subset P_n of all maps in S constructed in the following way : the unit cube $D = [0,1]^d$ is split into $N = 2^{nd}$ subcubes of size 2^{-n}, denoted by $D_{n,i}$, for $i = 1, ..., N$, with center of mass $x_{n,i}$. To each permutation σ of the N first integers, we associate the transform $\phi = \phi_\sigma$, $D \to D$, defined by

$$\phi(x) = x - x_{n,i} + x_{n,\sigma(i)}, \tag{55}$$

for all $x \in D_{n,i}$. Such a map will be called (with a slight abuse) a permutation. They form a set of $N!$ elements denoted by P_n and P will denote the collection of all P_n for $n \geq 0$, which clearly is a "subgroup" of the semi-group S. Apparently, G_0 and P are poorly related to each other. However, we claim

Proposition 64. *If $D = [0,1]^d$ with $d \geq 2$, then for all L^p norms $1 \leq p < +\infty$, P is contained in the closure of G_0.*

15.3.2 Proof of Proposition 64

Since every element of P can be written as a finite product of permutations exchanging adjacent subcubes (i.e. having a joint face), it is enough to show that such permutations can be approximated by a sequence in G_0, because of the following lemma :

Lemma 65. *Let S_1, S_2 two subsets of S contained in the closure of G_0 with respect to the L^p norm ($1 \leq p < +\infty$). Then this closure also contains $\{s_1 \circ s_2 \ ; \ s_1 \in S_1 \ , \ \ s_2 \in S_2\}$.*

Proof

Let $s_1 \in S_1$, $s_2 \in S_2$. For all g_1, g_2 in G_0,

$$||s_1 \circ s_2 - g_1 \circ s_2||_{L^p} = ||s_1 - g_1||_{L^p}$$

(since s_2 is MP (measure preserving)),

$$||g_1 \circ s_2 - g_1 \circ g_2||_{L^p} \leq Lip(g_1)||s_2 - g_2||_{L^p}$$

(since g_1 is Lipschitz continuous). Thus, we can make, by the triangle inequality,

$$||s_1 \circ s_2 - g_1 \circ g_2||_{L^p}$$

arbitrarily small by choosing first g_1 and, then, g_2, which completes the proof since $g_1 \circ g_2$ belongs to G_0.

To prove Proposition 64, it is now enough to approximate a permutation of two adjacent subcubes by a sequence in G_0. After obvious rescalings and translations, we are reduced to construct on the cube $Q =]-1,+1[\times]-1/2,1/2[^{d-1}$ a divergence free vector field

$$(t,x) \in]0,1[\times Q \to v(t,x) \in \mathbb{R}^d$$

smooth and compactly supported such that $g_1(v)$ is arbitrarily close (in L^p norm) to the map

$$(x_1, x_2, ..., x_d) \to (x_1 - sign(x_1), x_2, ..., x_d).$$

By using again the lemma, we can decompose this map and rather consider the (partial) symmetry map

$$(x_1, x_2, x_3, ..., x_d) \to (-x_1, -x_2, x_3, ..., x_d),$$

and two analogous maps on the cubes $Q_- =]-1,0[\times]-1/2,1/2[^{d-1}$ and $Q_+ =]0,+1[\times]-1/2,1/2[^{d-1}$. Let us only consider the first map. We introduce a so-called "stream function"

$$\psi(x_1, x_2) = \max(x_1^2, 4x_2^2) - 1.$$

and set

$$v(x) = (\partial_2\psi(x_1, x_2), -\partial_1\psi(x_1, x_2), 0, ..., 0).$$

Of course, this field is not smooth, but, we can already integrate it (because of its special structure, although the Cauchy-Lipschitz theorem does not apply) and get a non smooth flow $(t,x) \to g_t(v)(x)$ which exactly fits with our given symmetry map at time $t = 1$. (Exercise : compute all trajectories $dx/dt = v(x)$ in Q.) To get a smooth approximation $g_1(v_\epsilon) \in G_0$, it is enough to mollify v and rather consider $v_\epsilon \in V$ defined by

$$v_\epsilon(t,x) = \theta_\epsilon(t)(\partial_2\psi_\epsilon(x_1, x_2), -\partial_1\psi_\epsilon(x_1, x_2), 0, ..., 0),$$

where ψ_ϵ and θ_ϵ are suitable compactly supported smooth approximations of, respectively, ψ on $]-1,+1[\times]-1/2,+1/2[$ and 1 on $]0,1[$. (See more details for the mollification process in [198].) Notice that $d \geq 2$ is clearly needed to achieve the construction.

15.3.3 Bistochastic measures

To prove the density theorem it is now enough to show that P is a dense subset of S. As a matter of fact, we are going to prove a richer result based on the concept of "bistochastic measures" which are probabilistic generalizations of MPM (in the same way as Young's measures are generalization of functions in the framework of Calculus of Variations and non linear PDEs). For the definition, we go back, just for a short while, to a general setting.

Definition 66. *Let X and Y two topological spaces with Borel probability measures α and β, respectively. We say that a Borel probability measure μ on $X \times Y$ is bistochastic if its marginals are respectively α and β, namely*

$$\mu(A \times Y) = \alpha(A), \;\; \mu(X \times B) = \beta(B),$$

for all Borel subsets A and B of X and Y respectively.

This concept goes probably back to Kantorovich and was used to provide generalized solutions to the Monge optimal mass transfer problem, which will be discussed later in the course. There is a natural embedding of the set S of all MPP into the set DS of all bistochastic measures. Indeed, to each such map ϕ from (X, α) to (Y, β), we associate a unique μ in DS by setting

$$\mu(A \times B) = \alpha(\phi^{-1}(B) \cap A),$$

for all Borel subsets A and B of X and Y respectively, or, equivalently with distributional notations,

$$d\mu(x, y) = \delta(y - \phi(x))d\alpha(x),$$

where δ denotes the Dirac measure.

Exercise

Show that μ is bistochastic if and only if for all function f, $\alpha-$integrable on X, and for all function g, $\beta-$integrable on Y, $(x, y) \to (f(x), g(y))$ is $\mu-$integrable and

$$\int_{X \times Y} f(x)d\mu(x, y) = \int_X f(x)d\alpha(x),$$

$$\int_{X \times Y} g(y)d\mu(x, y) = \int_Y g(y)d\beta(y).$$

Exercise

Investigate the bistochastic measures as X and Y are finite set with discrete measures. Address, in particular, the case when $X = Y$ with the counting measure.

15.3.4 Density of P in S and DS

Let us now return to the case $X = Y = D = [0, 1]^d$ and $\alpha = \beta = |\cdot|$. To show the density of P in S, it is enough to show that P is densely embedded in DS, with respect to the vague topology of measures, thanks to the following lemma, which can be proved as an exercise.

Lemma 67. *Let (ϕ_n) a sequence in S and (μ_{ϕ_n}) the corresponding sequence in DS. Then ϕ_n converges to $\phi \in S$ for all L^p norm, $p < +\infty$, if and only if (μ_{ϕ_n}) vaguely converges to μ_ϕ.*

Thus we are left to show that for a fixed given $\mu \in DS$, there is a sequence of "permutations" (p_n) such that μ_{p_n} converges vaguely to μ.
Let $n > 0$ fixed integer and $N = n^d$. We split $D = [0,1]^d$ into N subcubes of equal volume denoted by $D_{n,i}$ for $i = 1, ..., N$. We set

$$\nu_{ij} = N\mu(D_{n,i} \times D_{n,j}),$$

for $i, j = 1, ..., N$ so that ν is a so-called $N \times N$, bistochastic matrix, i.e. a matrix with only nonnegative entries whose every column and every row add up to one. $>$From a classical result of G.Birkhoff, such a matrix always can be written as a convex combination of at most $K = K(N)$ (where $K(N) \leq CN^2$) permutation matrices. Thus, there are coefficients $\theta_1, ..., \theta_K \geq 0$ and permutations $\sigma_1, ..., \sigma_K$ such that

$$\sum_{k=1}^{K} \theta_k = 1, \quad \nu_{ij} = \sum_{k=1}^{K} \theta_k \delta_{j,\sigma_k(i)}.$$

Let us introduce $L = 2^{ld}$, where l will be chosen later, and set

$$\theta'_k = \frac{1}{L}([L\theta_k] + \epsilon_k),$$

where $[.]$ denotes the integer part of a real number and $\epsilon_k \in [0, 1[$ is chosen so that

$$\sum_{k=1}^{K} \theta'_k = 1, \quad \sup_k |\theta_k - \theta'_k| \leq \frac{1}{L}.$$

By setting

$$\nu'_{ij} = \sum_{k=1}^{K} \theta'_k \delta_{j,\sigma_k(i)},$$

we get a new bistochastic matrix which satisfies

$$\sum_{i,j} |\nu'_{ij} - \nu_{ij}| \leq \frac{NK}{L}.$$

Up to a relabelling of the list of permutations, with possible repetitions, we may assume all coefficients θ'_k to be equal to $1/L$ and get a new expression

$$\nu'_{ij} = \frac{1}{L} \sum_{k=1}^{L} \delta_{j,\sigma_k(i)}.$$

Now, we can split again each $D_{n,i}$ into L subcubes, denoted by $D_{n+l,i,m}$, for $i = 1, ..., N$, $m = 1, ..., L$, with size $2^{-(n+l)}$ and volume $2^{-(n+l)d}$. Then, we define

$$p(x) = x - x_{n+l,i,m} + x_{n+l,\sigma_m(i),m},$$

for each $x \in D_{n+l,i,m}$. By construction, $(i, m) \to (\sigma_m(I), m)$ is one-to-one. Thus, p belongs to P_{n+l}. Let us now estimate, for any fixed $f \in C(D)$,

$$I_1 - I_2 = \int_{D^2} f(x, y)\mu(dx, dy) - \int_D f(x, p(x))dx.$$

We denote by η the modulus of continuity of f. I_1 is equal, up to an error of $\eta(2^{-n+d/2})$, to

$$I_3 = \frac{1}{N} \sum_{i,j} f(x_{n,i}, x_{n,j}) \nu_{ij}.$$

I_3 is equal, up to an error of $\sup |f| K/L$ to

$$I_4 = \frac{1}{N} \sum_{i,j} f(x_{n,i}, x_{n,j}) \nu'_{ij} = \frac{1}{NL} \sum_{i,m} f(x_{n,i}, x_{n,\sigma_m(i)}).$$

Up to $\eta(2^{-n+d/2})$, I_4 is equal to

$$I_5 = \frac{1}{NL} \sum_{i,m} f(x_{n+l,i,m}, x_{n+l,\sigma_m(i),m}).$$

I_5, up to $\eta(2^{-n-l+d/2})$, is equal to

$$I_6 = \sum_{i,m} \int_{D_{n+l,i,m}} f(x, x - x_{n+l,i,m} + x_{n+l,\sigma_m(i),m}),$$

which is exactly I_2, by definition of p. Finally, we have shown

$$|I_1 - I_2| \leq \sup |f| 2^{(2n-l)d} + 3\eta(2^{-n-l+d/2}),$$

since $L = 2^{ld}$, $K = N^2 = 2^{2nd}$. This completes the proof, after letting first l and then n to $+\infty$.

15.3.5 Proof of the Birkhoff theorem

The proof relies on the classical "marriage lemma" from combinatorics, that asserts that a necessary and sufficient condition to marry N girls to N boys without dissatisfaction is that, for all subset of $r \leq N$ girls, there are at least r convenient boys. Let (ν_{ij}) be a bistochastic matrix. There is a permutation σ such that $\inf_i \nu_{i,\sigma(i)}$ is a positive number $\alpha > 0$. (In other words the "support" of σ is contained in the support of ν.) Then, we have the following alternative. Either $\alpha = 1$ and ν is automatically a permutation matrix. Or $\alpha < 1$ and

$$\nu'_{ij} = (\nu_{ij} - \alpha \delta_{j,\sigma(i)}) \frac{1}{1-\alpha}$$

defines a new bistochastic matrix with a strictly smaller support and ν is a convex combination of ν' and a permutation matrix. Recursively, after a finite number of steps, ν is written as a convex combination of permutation matrices which completes the proof.

15.4 Related density results

Using the marriage lemma, P. Lax has shown that if ϕ is a continuous MPM on $D = [0,1]^d$, then there is $p \in P_n$ such that

$$\sup_{x \in D} |\phi(x) - p(x)| \leq \eta(2^{-n+d/2}) + 2^{-n} C_d,$$

where η is the modulus of continuity of ϕ and C_d depends only on dimension d.

A sort of Lusin theorem holds true for MPM. More precisely, if ϕ is an almost everywhere one-to-one MPM, then, for all $\epsilon > 0$, there is a MPM homeomorphism (i.e. a one-to-one continuous MPM with continuous inverse) such that the measure of the set where ϕ differs from ϕ_ϵ is less than ϵ.

About this kind of questions, one may look at the books of Oxtoby (Springer lecture notes 318 (1973)) and Sudakov (Proc. Steklov Institute, 141, 1979).

References

[1] Y. Achdou, F. Camilli, I. Capuzzo-Dolcetta, *Mean field games: numerical methods for the planning problem,* SIAM J. Control Optim. 50 (2012) 77-109.

[2] J.-B. d'Alembert, *Recherches sur la courbe que forme une corde tendue mise en vibration,* Histoire de lâĂŹAcadémie royale des sciences et belles-lettres de Berlin pour lâĂŹannée 1747, 1750, 214-219.

[3] L. Ambrosio, *Transport equation and Cauchy problem for BV vector fields,* Invent. Math. 158 (2004) 227-260.

[4] L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli, *A global existence result for the semigeostrophic equations in three dimensional convex domains,* Discrete Contin. Dyn. Syst. 34 (2014) 1251-1268.

[5] L. Ambrosio, A. Figalli, *On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations,* Calc. Var. Partial Differential Equations 31 (2008) 497-509.

[6] L. Ambrosio, A. Figalli, *Geodesics in the space of measure-preserving maps and plans,* Archive for Rational Mechanics and Analysis 194 (2009) 421-462.

[7] L. Ambrosio, W. Gangbo, *Hamiltonian ODE in the Wasserstein spaces of probability measures,* Comm. Pure Appl. Math. 61(2008) 18-53.

[8] L. Ambrosio, N.Gigli, G. Savaré *Gradient flows in metric spaces and the Wasserstein spaces of probability measures,* Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.

[9] M. Arnaudon, A. B. Cruzeiro, N. Galamba, *Lagrangian Navier-Stokes flows: a stochastic model* J. Phys. A: Math. Theor. 44 (2011).

[10] M. Arnaudon, A. B. Cruzeiro, Ch. Léonard, J.-C. Zambrini *An entropic interpolation problem for incompressible viscid fluids* preprint arXiv:1704.02126.

[11] V. Arnold, *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications a l'hydrodynamique des fluides parfaits,* Ann. Institut Fourier 16 (1966) 319-361.

[12] V. Arnold, B. Khesin, *Topological methods in hydrodynamics, Applied Mathematical Sciences, 125,* Springer-Verlag 1998.

[13] J.-P. Aubin, *Mathematical methods of game and economic theory,* Studies in Mathematics and its Applications, 7. North-Holland 1979.

[14] V. Bach, J.-B. Bru, *Rigorous foundations of the Brockett-Wegner flow for operators,* J. Evol. Equ. 10 (2010) 425-442.

[15] M. Balinski, *A competitive (dual) simplex method for the assignment problem,* Math. Programming 34 (1986) 125-141.

[16] A. Baradat, *Continuous dependence of the pressure field with respect to endpoints for ideal incompressible fluids,* Calc. Var. Partial Differential Equations 58 (2019), no. 1, Art. 25, 22

[17] A. Baradat, *Nonlinear instability in Vlasov type equations around rough velocity profiles,* arXiv:1811.01350.

[18] A. Baradat, L. Monsaingeon, *Small noise limit and convexity for generalized incompressible flows, Schrödinger problems, and optimal transport,* arXiv:1810.12036.

[19] C. Bardos, N. Besse, *The Cauchy problem for the Vlasov-Dirac-Benney equation and related issues in fluid mechanics and semi-classical limits,* Kinet. Relat. Models 6 (2013) 893-917.

[20] C. Bardos, A.-Y. le Roux, J.-C. Nédélec, *First order quasilinear equations with boundary conditions,* Comm. Partial Differential Equations 4 (1979) 1017-1034.

[21] G. Bellettini, J. Hoppe, M. Novaga, G. Orlandi, *Closure and convexity results for closed relativistic strings,* Complex Anal. Oper. Theory 4 (2010) 473-496.

[22] J.-D. Benamou, Y. Brenier, *Weak existence for the semigeostrophic equation formulated as a coupled Monge-Ampere/transport problem,* SIAM J. Appl. Math., 58 (1998) 1450-1461.

[23] J.-D. Benamou, Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem,* Numer. Math. 84 (2000) 375-393.

[24] J.-D. Benamou, G. Carlier, L. Nenna, *Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm,* preprint, arXiv:1710.08234.

[25] J.-D. Benamou, M. Cuturi, G. Carlier, L. Nenna, *Iterative Bregman projections for regularized transportation problems*, SIAM J. Sci. Comput. 37 (2015) A1111-A1138.

[26] R. Berman, B. Berndtsson, *Real Monge-Ampère equations and Kähler-Ricci solitons on toric log Fano varieties,* Ann. Fac. Sci. Toulouse Math. (6) 22 (2013) 649-711.

[27] R. Berman, M. Önheim, *Propagation of chaos, Wasserstein gradient flows and toric Kähler-Einstein metrics,* Anal. PDE 11 (2018), no. 6, 1343-1380.

[28] R. Berman, M. Önheim, *Propagation of chaos for a class of first order models with singular mean field interactions,* SIAM J. Math. Anal. 51 (2019) 159-196.

[29] M. Bernot, A. Figalli, F. Santambrogio, *Generalized solutions for the Euler equations in one and two dimensions,* J. Math. Pures Appl. 91 (2009) 137-155.

[30] F. Berthelin, A. Vasseur, *From kinetic equations to multidimensional isentropic gas dynamics before shocks,* SIAM J. Math. Anal. 36 (2005) 1807-1835.

[31] S. Bianchini, A. Bressan, *Vanishing viscosity solutions of nonlinear hyperbolic systems,* Ann. of Math. (2) 161 (2005) 223-342.

[32] G. Boillat. C. Dafermos, P. Lax, T.P. Liu, *Recent mathematical methods in nonlinear wave propagation,* Lecture Notes in Math., 1640, Springer, Berlin, 1996.

[33] F. Bolley, Y. Brenier, G. Loeper, *Contractive metrics for scalar conservation laws,* J. Hyperbolic Differ. Equ. 2 (2005) 91-107.

[34] M. Born, L. Infeld, *Foundations of a new field theory*, Proc. Roy. Soc. London A 144 (1934) 425-451.

[35] F. Bouchut, *Renormalized solutions to the Vlasov equation with coefficients of bounded variation,* Arch. Ration. Mech. AnaBull. Inst. Math. Acad. Sin. (N.S.) 11 (2016), no. 1, 23?41.l. 157 (2001) 75-90.

[36] Y. Brenier, *Une application de la symétrisation de Steiner aux équations hyperboliques: la méthode de transport et écroulement,* C. R. Acad. Sci. Paris Ser. I Math. 292 (1981) 563-566.

[37] Y. Brenier, *Une équation homologique avec contrainte,* C. R. Acad. Sci. Paris Sér. I Math. 295 (1982) 103-106.

[38] Y. Brenier, *Résolution d'équations d'évolution quasilinéaires en dimension N d'espace à l'aide d'équations linéaires en dimension $N+1$*, J. Differential Equations 50 (1983) 375-390.

[39] Y. Brenier, *Averaged multivalued solutions for scalar conservation laws,* SIAM J. Numer. Anal. 21 (1984) 1013-1037.

[40] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs,* C. R. Acad. Sci. Paris I Math. 305 (1987) 805-808.

[41] Y. Brenier, *Un algorithme rapide pour le calcul de transformées de Legendre-Fenchel discretes,* C. R. Acad. Sci. Paris Sér I Math. 308 (1989) 587-589.

[42] Y. Brenier, *A combinatorial algorithm for the Euler equations of incompressible flows,* Comput. Methods Appl. Mech. Engrg. 75 (1989) 325-332.

[43] Y. Brenier, *The least action principle and the related concept of generalized flows for incompressible perfect fluids,* J.of the AMS 2 (1989) 225-255.

[44] Y. Brenier, *Une formulation de type Vlassov-Poisson pour les équations d'Euler des fluides parfaits incompressibles,* RR-107 INRIA, 1988. https://hal.inria.fr/inria-00075489

[45] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions,* Comm. Pure Appl. Math. 44 (1991) 375-417.

[46] Y. Brenier, *The dual least action principle for an ideal, incompressible fluid* Arch. Rational Mech. Anal. 122 (1993) 323-351.

[47] Y. Brenier, *A homogenized model for vortex sheets,* Arch. Rational Mech. Anal. 138 (1997) 319-353.

[48] Y. Brenier, *Minimal geodesics on groups of volume-preserving maps,* Comm. Pure Appl. Math. 52 (1999) 411-452.

[49] Y. Brenier, *Homogeneous hydrostatic flows with convex velocity profiles,* Nonlinearity 12 (1999) 495-512.

[50] Y. Brenier, *Derivation of the Euler equations from a caricature of Coulomb interaction,* Comm. Math. Phys. 212 (2000) 93-104.

[51] Y. Brenier, *Convergence of the Vlasov-Poisson system to the incompressible Euler equations*, Comm. Partial Differential Equations 25 (3-4) (2000) 737-754.

[52] Y. Brenier, *Order preserving vibrating strings and applications to electrodynamics and magnetohydrodynamics*, Methods Appl. Anal. 11 (2004) 515-532.

[53] Y. Brenier, *Remarks on the derivation of the hydrostatic Euler equations,* Bull. Sci. Math. 127 (2003) 585-595.

[54] Y. Brenier, *Extended Monge-Kantorovich theory, Optimal transportation and applications,* pp. 91-121, Lecture Notes in Math., 1813, Springer 2003.

[55] Y. Brenier, *Hydrodynamic structure of the augmented Born-Infeld equations,* Arch. Ration. Mech. Anal. 172 (2004) 65-91.

[56] Y. Brenier, *Non relativistic strings may be approximated by relativistic strings,* Methods Appl. Anal. 12 (2005) 153-167.

[57] Y. Brenier, *Generalized solutions and hydrostatic approximation of the Euler equation,* Phys. D 237 (2008) 14-17.

[58] Y. Brenier, *L2 formulation of multidimensional scalar conservation laws,* Arch. Ration. Mech. Anal. 193 (2009) 1-19.

[59] Y. Brenier, *Optimal transport, convection, magnetic relaxation and generalized Boussinesq equations*, J. Nonlinear Sci. 19 (2009) 547-570.

[60] Y. Brenier, *On the hydrostatic and Darcy limits of the convective Navier-Stokes equations*, Chin. Ann. Math. Ser. B 30 (2009) 683-696.

[61] Y. Brenier, *A modified least action principle allowing mass concentrations for the early universe reconstruction problem,* Confluentes Mathematici 3 (2011) 361-385.

[62] Y. Brenier, *Remarks on the Minimizing Geodesic Problem in Inviscid Incompressible Fluid Mechanics,* Calc. Var. Partial Differential Equations 47 (2013) 55-64.

[63] Y. Brenier, *Rearrangement, convection, convexity and entropy,* Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 371 (2013), no. 2005, 20120343.

[64] Y. Brenier, *Topology-preserving diffusion of divergence-free vector fields and magnetic relaxation,* Comm. Math. Physics 330 (2014) 757-770.

[65] Y. Brenier, *Connections between optimal transport, combinatorial optimization and hydrodynamics,* ESAIM Math. Model. Numer. Anal. 49 (2015) 1593-1605.

[66] Y. Brenier, *A double large deviation principle for Monge-Ampère gravitation,* Bull. Inst. Math. Acad. Sin. (N.S.) 11 (2016) 23-41.

[67] Y. Brenier, *The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem,* Comm. Math. Phys. 364 (2018) 579-605.

[68] Y. Brenier, *Geometric origin and some properties of the arctangential heat equation,* Tunis. J. Math. 1 (2019) 561-584.

[69] Y. Brenier, L. Corrias, *A kinetic formulation for multi-branch entropy solutions of scalar conservation laws,* Ann. Inst. H. Poincaré Anal. Non Linéaire 15 (1998) 169-190.

[70] Y. Brenier, C. De Lellis, L. Székelyhidi Jr., *Weak-strong uniqueness for measure-valued solutions,* Comm. Math. Phys. 305 (2011) 351-361.

[71] Y. Brenier, X. Duan, $>$*From Conservative to Dissipative Systems Through Quadratic Change of Time, with Application to the Curve-Shortening Flow,* Arch. Rational Mech. Anal. 227 (2018) 545-565.

[72] Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, Mohayaee, Sobolevskii, *Reconstruction of the early universe as a convex optimization problem,* Mon. Not. R. Astron. Soc. 2002.

[73] Y. Brenier, W. Gangbo, L^p *approximation of maps by diffeomorphisms,* Calc. Var. Partial Differential Equations 16 (2003) 147-164.

[74] Y. Brenier, W. Gangbo, G. Savaré, M. Westdickenberg, *Sticky particle dynamics with interactions* J. Math. Pures Appl. (9) 99 (2013) 577–617

[75] Y. Brenier, E. Grenier, *Sticky particles and scalar conservation laws,* SIAM J. Numer. Anal. 35 (1998) 2317-2328.

[76] Y. Brenier, J. Jaffré, *Upstream differencing for multiphase flow in reservoir simulation,* SIAM J. Numer. Anal. 28 (1991) 685-696.

[77] Y. Brenier, G., Loeper, *A geometric approximation to the Euler equations: The Vlasov-Monge- Ampère equation,* Geom. Funct. Anal. 14(2004) 1182-1218.

[78] Y. Brenier, M. Puel, *Optimal multiphase transportation with prescribed momentum,* ESAIM Control Optim. Calc. Var. 8 (2002) 287-343.

[79] Y. Brenier, W. Yong, *Derivation of particle, string, and membrane motions from the Born-Infeld electromagnetism,* J. Math. Phys. 46 (2005), no. 6, 062305.

[80] P. Brenner, *The Cauchy problem for symmetric hyperbolic systems in* L_p, Math. Scand. 19 (1966) 27-37.

[81] D. Bresch, A. Kazhikhov, J. Lemoine, *On the two-dimensional hydrostatic Navier-Stokes equations,* SIAM J. Math. Anal. 36 (2004/05) 796-814.

[82] A. Bressan, *Hyperbolic systems of conservation laws. The one-dimensional Cauchy problem,* Oxford University Press, 2000.

[83] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert,* North-Holland Mathematics Studies, No. 5. 1973.

[84] H. Brezis, F. Browder, *Partial differential equations in the 20th century,* Adv. Math. 135 (1998) 1, 76-144.

[85] H. Brezis, M. Crandall, *Uniqueness of solutions of the initial-value problem for $u_t - \Delta\phi(u) = 0$,* J. Math. Pures Appl. (9) 58 (1979) 153-163.

[86] T. Buckmaster, C. de Lellis, L. Székelyhidi Jr, V. Vicol, *Onsager's conjecture for admissible weak solutions,* Comm. Pure Appl. Math. 72 (2019) 229-274.

[87] G. R. Burton, *Rearrangements of functions, maximization of convex functionals and vortex rings,* Math. Ann., 276, 225-253, 1987

[88] L. Caffarelli *The regularity of mappings with a convex potential,* J. Amer. Math. Soc. 5 (1992) 99-104.

[89] P. Cardaliaguet, F. Delarue, J.-M. Lasry, P.-L. Lions, *The Master Equation and the Convergence Problem in Mean Field Games,* Annals of Math. Studies 201, Princeton University Press, 2019.

[90] J. A. Carrillo, M. Di Francesco, C. Lattanzio, *Contractivity of Wasserstein Metrics and Asymptotic Profiles for Scalar Conservation Laws,* J. Differential Equations 231 (2006) 425-458.

[91] N. Champagnat, P.-E. Jabin, *Well posedness in any dimension for Hamiltonian flows with non BV force terms,* Comm. Partial Differential Equations 35 (2010) 786-816.

[92] G. Chavent, J. Jaffré, *Mathematical Models and Finite Elements for Reservoir Simulation,* North Holland, 2000.

[93] J.-Y. Chemin, *Existence globale pour le problème des poches de tourbillon,* C. R. Acad. Sci. Paris Sér. I Math. 312 (1991) 803-806.

[94] J.-Y. Chemin,

[95] J.-Y. Chemin, *Perfect incompressible fluids,* Oxford University Press, 1998.

[96] J.-Y. Chemin, B. Desjardins, I. Gallagher, E. Grenier, *Mathematical geophysics. An introduction to rotating fluids and the Navier-Stokes equations,* Oxford University Press 2006.

[97] L. Chizat, G. Peyré, B. Schmitzer, F.-X. Vialard, *Unbalanced Optimal Transport: Geometry and Kantorovich Formulation,* J. Funct. Anal. 274 (2018) 3090-3123.

[98] P. Constantin, Weinan E, E. Titi, *Onsager's conjecture on the energy conservation for solutions of Euler's equation,* Comm. Math. Phys. 165 (1994) 207-209.

[99] L. Corrias, *Fast Legendre-Fenchel transform and applications to Hamilton-Jacobi equations and conservation laws,* SIAM J. Numer. Anal. 33 (1996) 1534-1558.

[100] D. Christodoulou, S. Klainerman, *The global nonlinear stability of the Minkowski space,* Princeton Mathematical Series, 41, Princeton University Press 1993.

[101] D. Cordero-Erausquin, B. Nazaret, C. Villani, *A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities,* Adv. Math. 182 (2004), no. 2, 307âĂŞ332.

[102] M. Crandall, L.C. Evans, P.-L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations,* Trans. Amer. Math. Soc. 282 (1984) 487-502.

[103] G. Crippa, C. De Lellis, *Estimates and regularity results for the DiPerna-Lions flow,* J. Reine Angew. Math. 616 (2008) 15-46.

[104] M. Cullen, W. Gangbo, L. Pisante, *The semigeostrophic equations discretized in reference and dual variables,* Arch. Ration. Mech. Anal.,1185 (2007) 341-363.

[105] M.Cullen, J. Purser, *An extended Lagrangian theory of semigeostrophic frontogenesis,* J. Atmospheric Sci. 41 (1984) 1477-1497.

[106] C. Dafermos, *Hyperbolic conservation laws in continuum physics, Fourth edition,* Springer-Verlag, Berlin, 2016.

[107] C. De Lellis, L. Székelyhidi Jr, *The Euler equations as a differential inclusion,* Ann. of Math. (2) 170 (2009) 1417-1436.

[108] C. De Lellis, L. Székelyhidi Jr, *On admissibility criteria for weak solutions of the Euler equations,* Arch. Ration. Mech. Anal. 195 (2010) 225-260.

[109] C. De Lellis, L. Székelyhidi Jr, *On turbulence and geometry: from Nash to Onsager,* Notices Amer. Math. Soc. 66 (2019) 677-685.

[110] J.-M. Delort, *Existence de nappes de tourbillon en dimension deux,* J. Amer. Math. Soc. 4 (1991) 553-586.

[111] G. De Philippis, A. Figalli, *W2,1 regularity for solutions of the Monge-Ampere equation,* Invent. Math. 192 (2013) 55-69.

[112] G. De Philippis, A. Figalli, *The Monge-AmpÃČÅąre equation and its link to optimal transportation,* Bull. Amer. Math. Soc. (N.S.) 51 (2014) 527-580.

[113] B. Després, C. Mazeran, *Constant Lagrangian gas dynamics in two dimensions and Lagrangian systems,* Arch. Ration. Mech. Anal. 178 (2005) 327-372.

[114] R.J. DiPerna, *Convergence of the viscosity method for isentropic gas dynamics,* Comm. Math. Phys. 91 (1983) 1-30.

[115] R.J. DiPerna, *Measure-valued solutions to conservation laws,* Arch. Rational Mech. Anal. 88 (1985) 223-270.

[116] R.J. DiPerna, *Compensated compactness and general systems of conservation laws,* Trans. Amer. Math. Soc. 292 (1985) 383-420.

[117] R.J. DiPerna, P.-L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces,* Invent. Math. 98 (1989) 511-547.

[118] R.J. DiPerna, A. Majda *Oscillations and concentrations in weak solutions of the incompressible fluid equations,* Comm. Math. Phys. 108 (1987) 667-689.

[119] X. Duan, *Hyperbolicity of the time-like extremal surfaces in Minkowski spaces,* arXiv:1706.04768.

[120] J. Duchon, R. Robert, *Relaxation of the Euler equations and hydrodynamic instabilities,* Quarterly Appl. Math. 50 (1992) 235-255.

[121] D. Ebin, J. Marsden, *Groups of diffeomorphisms and the notion of an incompressible fluid,* Ann. of Math. 92 (1970) 102-163.

[122] M. Edelstein *On nearest points of sets in uniformly convex Banach spaces,* J. London Math. Soc. 43 (1968) 375-377.

[123] I. Ekeland, *The Hopf-Rinow theorem in infinite dimension,* J. Diff. Geo. 13 (1978) 287-301.

[124] Y.Eliashberg, T.Ratiu, *The diameter of the symplectomorphism group is infinite,* Invent. Math. 103 (1991) 327-340.

[125] B. Engquist, B. Froese, Y. Yang, *Optimal transport for seismic full waveform inversion,* Commun. Math. Sci. 14 (2016) 2309-2330.

[126] B. Engquist, O. Runborg, *Computational high frequency wave propagation,* Acta Numer. 12 (2003) 181-266.

[127] L. Euler, *Opera Omnia, Series Secunda, 12, 274-361.*

[128] L.C. Evans, EDP *Partial differential equations,* American Mathematical Society, 2010.

[129] L.C. Evans, W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem,* Mem. Amer. Math. Soc. 137 (1999).

[130] L.C. Evans, R. F. Gariepy, *Measure theory and fine properties of functions,* Revised edition. Textbooks in Mathematics, CRC Press, Boca Raton 2015.

[131] G. L. Eyink, *Energy dissipation without viscosity in ideal hydrodynamics,* Phys. D 78 (1994) 222-240.

[132] H. Federer, *Geometric measure theory,* Springer-Verlag 1969.

[133] R. Fedkiw, G. Sapiro, C-W Shu, *shock capturing, level sets, and PDE based methods in computer vision and image processing: a review of Osher's contributions,* J. Comput. Phys. 185 (2003) 309-341.

[134] E. Feireisl, P. Gwiazda, A. Świerczewska-Gwiazda, E. Wiedemann, *Dissipative measure-valued solutions to the compressible Navier-Stokes system,* Calc. Var. Partial Differential Equations (2016) 55:141, 20 pp.

[135] R. Feynman, *Le cours de physique de Feynman - Electromagnétisme 2 - 2e édition,* Dunod 2017.

[136] A. Figalli, L. Rifford, *Mass transportation on sub-Riemannian manifolds,* Geom. Funct. Anal. 20 (2010) 124-159.

[137] U. Frisch, *Turbulence. The legacy of A. N. Kolmogorov,* Cambridge University Press, 1995.

[138] U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, *reconstruction of the initial conditions of the Universe by optimal mass transportation,* Nature 417 (2002) 260-262.

[139] T. Gallouët, Q. Mérigot, *A Lagrangian scheme a la Brenier for the incompressible Euler equations,* Found. Comput. Math. 18 (2018) 835-865.

[140] D. Gérard-Varet, N. Masmoudi, *Well-posedness for the Prandtl system without analyticity or monotonicity,* Ann. Sci. Ec. Norm. Supér. (4) 48 (2015) 1273-1325.

[141] B. Gess, B. Perthame, P. Souganidis, *Semi-discretization for stochastic scalar conservation laws with multiple rough fluxes,* SIAM J. Numer. Anal. 54 (2016) 2187-2209.

[142] B. Gess, J. Sauer, E. Tadmor, *Optimal regularity in time and space for the porous medium equation,* arXiv:1902.08632

[143] N. Ghoussoub, *Self-dual partial differential systems and their variational principles,* Springer Monographs in Mathematics, 2009.

[144] N. Ghoussoub, A. Moameni, *Symmetric Monge-Kantorovich problems and polar decompositions of vector fields,* Geom. Funct. Anal. 24 (2014) 1129-1166.

[145] U. Gianazza, G. Savaré, G. Toscani, *The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation,* Arch. Ration. Mech. Anal. 194 (2009) 133-220.

[146] J. Giesselmann, C. Lattanzio, A. Tzavaras, *Relative energy for the Korteweg theory and related Hamiltonian flows in gas dynamics*, Arch. Ration. Mech. Anal. 223 (2017) 1427-1484.

[147] M-H. Giga, Y. Giga, *Minimal vertical singular diffusion preventing overturning for the Burgers equation,* Recent advances in scientific computing and PDEs, Contemp. Math., 330, Amer. Math. Soc., 2003.

[148] Y. Giga, T. Miyakawa, *A kinetic construction of global solutions of first order quasilinear equations*, Duke Math. J. 50 (1983) 505-515.

[149] N. Gigli, F. Otto, *Entropic Burgers' equation via a minimizing movement scheme based on the Wasserstein metric*, Calc. Var. Partial Differential Equations 47 (2013) 181-206.

[150] M. Goldman, F Otto, *A variational proof of partial regularity for optimal transportation maps,* arXiv:1704.05339

[151] F. Golse, P.-L. Lions, B. Perthame, R. Sentis, *Regularity of the moments of the solution of a transport equation*, J. Funct. Anal. 76 (1988) 110-125.

[152] F. Golse, L. Saint-Raymond, *The Vlasov-Poisson system with strong magnetic field in quasineutral regime*, Math. Models Methods Appl. Sci. 13 (2003) 661-714.

[153] L. Gosse, G. Toscani, *Identification of asymptotic decay to self-similarity for one-dimensional filtration equations*, SIAM J. Numer. Anal. 43 (2006) 2590-2606.

[154] E. Grenier, *Defect measures of the Vlasov-Poisson system in the quasineutral regime,* Comm. Partial Differential Equations 20 (1995) 1189-1215.

[155] E. Grenier, *On the derivation of homogeneous hydrostatic equations,* M2AN Math. Model. Numer. Anal. 33 (1999) 965-970.

[156] M. Gromov, *Partial Differential Relations,* Springer-Verlag 1986.

[157] D. Han-Kwan, M. Iacobelli, *Quasineutral limit for Vlasov-Poisson via Wasserstein stability estimates in higher dimension,* J. Differential Equations 263 (2017) 1-25.

[158] D. Han-Kwan, F. Rousset, *Quasineutral limit for Vlasov-Poisson with Penrose stable data,* Ann. Sci. ENS (4) 49 (2016) 1445-1495.

[159] L. Hörmander, *The analysis of linear partial differential operators. III,* Springer-Verlag 1985.

[160] Ph. Isett, *A proof of Onsager's conjecture,* Ann. of Math. (2) 188 (2018) 871-963.

[161] S. Klainerman, *The null condition and global existence to nonlinear wave equations,* Nonlinear systems of partial differential equations in applied mathematics, 293-326, Lectures in Appl. Math., 23, Amer. Math. Soc., 1986.

[162] S. Klainerman, *PDE as a unified subject,* Geom. Funct. Anal. 2000, Special Volume, Part I, 279-315.

[163] S. Klainerman, A. Majda, *Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids,* Comm. Pure Appl. Math. 34 (1981) 481-524.

[164] Knott, Smith *On the optimal mapping of distributions,* J. Optim. Theory Appl. 43 (1984) 39-49.

[165] H. O. Kreiss, *Problems with different time scales for partial differential equations,* Comm. Pure Appl. Math. 33 (1980) 399-439.

[166] S. N. Kruzhkov, *First order quasilinear equations with several independent variables,* Mat. Sb. (N.S.) 81 (123) (1970) 228-255.

[167] J.-M. Lasry, P.-L. Lions, *Mean field games,* Jpn. J. Math. 2 (2007) 229-260.

[168] C. Lattanzio, A. Tzavaras, >*From gas dynamics with large friction to gradient flows describing diffusion theories,* Comm. Partial Differential Equations 42 (2017) 261-290.

[169] H. Lavenant, *Time-convexity of the entropy in the multiphasic formulation of the incompressible Euler equation,* Calc. Var. Partial Differential Equations 56 (2017) Art. 170, 29 pp.

[170] Ch. Léonard, *A survey of the Schrödinger problem and some of its connections with optimal transport,* Discrete Contin. Dyn. Syst. A, 34 (2014) 1533-1574

[171] J. Leray, *Sur le mouvement d'un liquide visqueux emplissant l'espace,* Acta Math. 63 (1934) 193-248.

[172] M. Liero, A. Mielke, G. Savaré, *Optimal Entropy-Transport problems and a new Hellinger-Kantorovich distance,* arXiv:1508.07941.

[173] H. Lindblad, *A remark on global existence for small initial data of the minimal surface equation in Minkowskian space time*, Proc. Amer. Math. Soc. 132 (2004) 1095-110.

[174] P.-L. Lions, *Mathematical topics in fluid mechanics. Vol. 1. Incompressible models,* Oxford Lecture Series in Mathematics and its Applications, 3. 1996.

[175] P.-L. Lions, ????????

[176] P.-L. Lions, S. Mas-Gallic, *Une méthode particulaire déterministe pour des équations diffusives non linéaires,* C. R. Acad. Sci. Paris Sér. I Math. 332 (2001) 369-376.

[177] P.-L. Lions, B. Perthame, P. Souganidis, *Existence and stability of entropy solutions for the hyperbolic systems of isentropic gas dynamics in Eulerian and Lagrangian coordinates,* Comm. Pure Appl. Math. 49 (1996) 599-638.

[178] P.-L. Lions, B. Perthame, E. Tadmor, *A kinetic formulation of multidimensional scalar conservation laws and related equations,* J. Amer. Math. Soc. 7 (1994) 169-191.

[179] G. Loeper, *Quasi-neutral limit of the Euler-Poisson and Euler-Monge-Ampere systems,* Comm. Partial Differential Equations 30 (2005) 1141-1167.

[180] G. Loeper, *Uniqueness of the solution to the Vlasov-Poisson system with bounded density,* J. Math. Pures Appl. (9) 86 (2006) 68-79.

[181] G. Loeper, *The reconstruction problem for the Euler-Poisson system in cosmology,* Arch. Ration. Mech. Anal. 179 (2006) 153-216.

[182] G. Loeper, *On the regularity of solutions of optimal transportation problems,* Acta Math. 202 (2009) 241-283.

[183] J. Lott, C. Villani, *Ricci curvature for metric-measure spaces via optimal transport,* Ann. of Math. (2) 169 (2009) 903-991.

[184] Y. Lucet, *Faster than the fast Legendre transform, the linear-time Legendre transform,* Numer. Algorithms, 16(2) (1998) 171-185.

[185] R. McCann, *A convexity principle for interacting gases,* Adv. Math. 128 (1997) 153-179.

[186] R. McCann, *Polar factorization of maps on Riemannian manifolds,* Geom. Funct. Anal. 11 (2001) 589-608.

[187] A. Majda, *Compressible fluid flow and systems of conservation laws in several space variables,* Applied Mathematical Sciences, 53. Springer-Verlag, 1984.

[188] A. Majda, A. Bertozzi, *Vorticity and incompressible flow,* Cambridge University Press, 2002.

[189] C. Marchioro, M. Pulvirenti *Mathematical theory of incompressible non viscous fluids,* Springer-Verlag, 1994.

[190] J. Marsden, T. Ratiu, *Introduction to Mechanics and Symmetry,* Texts in Applied Math. 17, Springer 1999.

[191] N. Masmoudi, T.K. Wong, *On the H^s theory of hydrostatic Euler equations,* Arch. Ration. Mech. Anal. 204 (2012) 231-271.

[192] Q. Mérigot, *A multiscale approach to optimal transport,* Computer Graphics Forum (2011) 30(5):1583-1592, 2011.

[193] Q. Mérigot, J.-M. Mirebeau, *Minimal geodesics along volume preserving maps, through semi-discrete optimal transport,* SIAM J. Numer. Anal. 54 (2016) 3465-3492.

[194] F. Morgan, Geometric measure theory. A beginner's guide. Academic Press 2000.

[195] S. Müller, M. Palombaro, *On a differential inclusion related to the Born-Infeld equations,* SIAM J. Math. Anal. 46 (2014) 2385-2403.

[196] F. Murat, *Compacité par compensation,* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 5 (1978) 489-507.

[197] J. Nash, *C1 isometric imbeddings,* Ann. of Math. (2) 60 (1954) 383-396.

[198] Y. Neretin, *Categories of bistochastic measures and representations of some infinite-dimensional groups,* Sb. 183 (1992), no. 2, 52-76.

[199] S. Osher, R. Fedkiw, *Level set methods. Geometric level set methods in imaging, vision, and graphics,* Springer, New York, 2003.

[200] F. Otto, *Evolution of microstructure in unstable porous media flow: a relaxational approach,* Comm. Pure Appl. Math. 52 (1999) 873-915.

[201] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation,* Comm. Partial Differential Equations 26 (2001) 101-174.

[202] F. Otto, M. Westdickenberg, *Eulerian calculus for the contraction in the Wasserstein distance,* SIAM J. Math. Anal. 37 (2005) 1227-1255.

[203] E. Y. Panov, *On kinetic formulation of first-order hyperbolic quasilinear systems,* Ukrainian Math. Vistnik 1(4) (2004) 548-563.

[204] M. Perepelitsa, *A note on strong solutions to the variational kinetic equation for scalar conservation laws,* Journal of Hyperbolic Differential Equations 11 (2014) 621-632.

[205] J. Polchinski, *String theory. Vol. I.* Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1998.

[206] M. Puel, L. Saint-Raymond, *Quasineutral limit for the relativistic Vlasov-Maxwell system,* Asymptot. Anal. 40 (2004) 303-352.

[207] T. Qin, *Symmetrizing nonlinear elastodynamic system,* J. Elasticity 50 (1998) 245-252.

[208] S.T.Rachev, L. Rüschendorf, *Mass transportation problems. Vol. I. and II.,* Springer-Verlag, 1998.

[209] L. Saint-Raymond, *Convergence of solutions to the Boltzmann equation in the incompressible Euler limit,* Arch. Ration. Mech. Anal. 166 (2003) 47-80.

[210] F. Santambrogio, *Optimal transport for applied mathematicians,* Birkhäuser/Springer, 2015.

[211] V. Scheffer, *An inviscid flow with compact support in space-time,* J. Geom. Anal. 3 (1993) 343-401.

[212] D. Serre, *Systems of conservation laws,* Cambridge University Press, Cambridge, 2000.

[213] D. Serre, *Hyperbolicity of the nonlinear models of Maxwell's equations,* Arch. Ration. Mech. Anal. 172 (2004) 309-331.

[214] D. Serre, *Multidimensional shock interaction for a Chaplygin gas,* Arch. Ration. Mech. Anal. 191 (2009) 539-577.

[215] D. Serre, *About the Young measures associated with Y. Brenier's ABI model,* J. Differential Equations 256 (2014) 3709-3720.

[216] D. Serre, A. Vasseur, *L2-type contraction for systems of conservation laws,* J. Ec. polytech. Math. 1 (2014) 1-28.

[217] S. Shandarin, Y. Zeldovich, *The large-scale structure of the universe: turbulence, intermittency, structures in a self-gravitating medium,* Rev. Modern Phys. 61 (1989) 185-220.

[218] A. Shnirelman, *On the principle of the shortest way in the dynamics of systems with constraints,* Global analysis studies and applications, II, 117-130, Lecture Notes in Math 1214, Springer, 1986.

[219] A. Shnirelman, *On the geometry of the group of diffeomorphisms and the dynamics of an ideal incompressible fluid,* Math. Sbornik USSR 56 (1987) 79-105.

[220] A. Shnirelman, *Generalized fluid flows, their approximation and applications,* Geom. Funct. Anal. 4 (1994) 586-620.

[221] A. Shnirelman, *On the nonuniqueness of weak solution of the Euler equation,* Comm. Pure Appl. Math. 50 (1997) 1261-1286.

[222] J. Speck, *The nonlinear stability of the trivial solution to the Maxwell-Born-Infeld system,* J. Math. Phys. 53 (2012), no. 8, 083703, 83 pp.

[223] K.-T. Sturm, *Convex functionals of probability measures and nonlinear diffusions on manifolds,* J. Math. Pures Appl. (9) 84 (2005) 149-168.

[224] L. Székelyhidi, Relaxation of the incompressible porous media equation, Ann. Sci. Ec. Norm. Supér. (4) 45 (2012) 491-509.

[225] L. Tartar, *Compacité par compensation: résultats et perspectives,* Nonlinear partial differential equations and their applications, Res. Notes in Math., 84, Pitman, Boston 1983.

[226] Taylor

[227] Q. H. Tran, M. Baudin, F. Coquel, *A relaxation method via the Born-Infeld system,* Math. Models Methods Appl. Sci. 19 (2009) 1203-1240.

[228] Y-H. R. Tsai, Y. Giga, S. Osher, *A level set approach for computing discontinuous solutions of Hamilton-Jacobi equations,* Math. Comp. 72 (2003) 159-181.

[229] J.L. Vázquez, *The Porous Medium Equation,* Oxford Univ. Press, 2007.

[230] C. Villani, *Topics in optimal transportation,* Graduate Studies in Mathematics, 58, AMS, Providence, 2003.

[231] C. Villani, *Optimal Transport, Old and New,* Springer-Verlag, 2009.

[232] E. Wiedemann, *Existence of weak solutions for the incompressible Euler equations,* Ann. Inst. H. Poincaré Anal. Non Linéaire 28 (2011) 727-730.

[233] G. Wolansky, *On time reversible description of the process of coagulation and fragmentation,* Arch. Ration. Mech. Anal. 193 (2009) 57-115.

[234] W. Wolibner, *Un théorème sur l'existence du mouvement plan d'un fluide parfait, homogène, incompressible,* Math. Z. 37(1) (1933) 698-726.

[235] L.C.Young, *Lectures on the calculus of variations,* Chelsea,New York, 1980.

[236] V. Yudovich, *Non-stationary flows of an ideal incompressible fluid,* Zh. Vychisl. Mat i Mat. Fiz. 3 (1963) 1032-1066.

[237] Y. Zeldovich, *Gravitational instability: An approximate theory for large density perturbations,* Astron. Astrophys. 5, 84-89 (1970).