TOPICS ON HYDRODYNAMICS AND VOLUME PRESERVING MAPS

*Yann Brenier**

Abstract

Various topics related to the mathematical theory of ideal incompressible fluids are discussed in relation with the concept of measure preserving maps.

1 General Presentation

1.1 The configuration space of an incompressible fluid

In Classical Continuum Mechanics [Ar], the motion of an incompressible fluid moving in a compact domain D of the Euclidean space $\mathbf{R}^d$ can be seen as a trajectory $t \rightarrow g(t)$ on the configuration space $G(D)$ (denoted by $SDiff(D)$ in [Ar]) of all diffeomorphisms of D with unit jacobian determinant. This configuration space can be embedded in a larger one, namely the set $S = S(D)$ of all maps h from D into itself, not necessarily one-to-one, such that, for all Borel subset B of D, $h^{-1}(B)$ is a Borel subset of D having the same Lebesgue measure as B. Equivalently, we can say that a Borel map h belongs to $S(D)$ if

$$\int_D \phi(h(x))dx = \int_D \phi(x)dx$$

holds true for all $\phi \in C(D)$, where dx denotes the Lebesgue measure, normalized so that the measure of D is 1, and $C(D)$ denotes the Banach space of all real continuous functions on D. For the composition rule, $G(D)$ is a group (the identity map I being the unity of the group), meanwhile $S(D)$ is a semi-group. Both $G(D)$ and $S(D)$ are naturally embedded in the Hilbert

*Institut Universitaire de France, CNRS UMR 6621, Nice, en détachement de l'Université Paris 6, France, brenier@math.unice.fr

space $L^2(D, \mathbf{R}^d)$ of all square integrable mapping from D into $\mathbf{R}^d$. One of our goals is to compare $S(D)$ and $G(D)$. It will be shown, in the simple case $D = [0,1]^d$, that $S(D)$ is closed in L^2 with $G(D)$ as a dense subset, for all $d \geq 2$. In the following discussions, it will be always assumed that D is chosen so that $S(D)$ is the L^2 closure of $G(D)$.

1.2 The Euler equations of incompressible fluids

An ideal incompressible fluid moving inside D is usually described by a velocity field $v(t,x)$ and a pressure field $p(t,x)$, subject to the classical Euler equations [Eu]

$$\partial_t v + (v.\nabla)v = -\nabla p,$$

$$\nabla.v = 0,$$

with the boundary condition that v is parallel to ∂D. Here we use standard PDE and Continuum Mechanics notation,

$$\nabla = (\partial_{x_1}, ..., \partial_{x_d})$$

and . denotes the inner product in the Euclidean space.
The flow $(t,x) \to g(t,x)$ describing the motion of fluid particles is defined by

$$\partial_t g(t,x) = v(t, g(t,x)), \quad g(0,x) = x,$$

and can be solved by the Cauchy-Lipschitz theorem provided that v is smooth enough and parallel to D. By elementary calculations, the Euler equations can be replaced by the following equivalent set of equations

$$\partial^2_{tt} g(t,x) = -(\nabla p)(t, g(t,x)), \quad \det(\partial_x g(t,x)) = 1, \tag{1}$$

which insures that $t \to g(t)$ is valued in the configuration space $G(D)$, provided that v is smooth enough. These equations have been established by Euler (Opera Omnia, Series secunda, livre 12, pp. 274-315 et 316-361), and still receive a lot of attention from Mathematicians, mostly but not only in the field of nonlinear PDEs, as shown, for instance, by the recent publication of various books by Arnold-Khesin, Chemin, P.-L. Lions, Marchioro-Pulvirenti etc... or Majda's conference at the ICM of Kyoto [Ma]. The first attempts to address the Euler equations as a nonlinear PDE problem (existence, uniqueness, stability of solutions) go back to the 20' [Lic].

1.3 Geometric interpretation of the Euler equations

A formal, non rigorous, Riemannian structure is naturally induced on the configuration space $G(D)$ by the L^2 norm of the ambient Hilbert space $L^2(D, \mathbf{R}^d)$ of all square integrable mapping from D into $\mathbf{R}^d$. The L^2 norm is defined by

$$||h||_{L^2}^2 = \int_D |h(x)|^2 dx,$$

where $|.|$ is the Euclidean norm.

We may *formally* define, for any pair g_0, g_1 in $G(D)$, their geodesic distance

$$\delta_D(g_0, g_1) = \inf \int_0^1 ||\partial_t g(t,.)||_{L^2} dt$$

where the infimum is performed over all smooth trajectories $t \to g(t) \in G(D)$ satisfying

$$g(0) = g_0, \quad g(1) = g_1.$$

A geodesic curve can be defined as a curve $t \to g(t) \in G(D)$ such that for all $t_0 \in \mathbf{R}$, there is $\delta > 0$ such that if $t_0 < t_1 < t_0 + \delta$, then

$$\delta_D(g(t_0), g(t_1)) = \int_{t_0}^{t_1} ||\partial_t g(t,.)||_{L^2} dt. \tag{2}$$

If, in addition, the t parametrization of g is chosen so that $||\partial_t g(t,.)||_{L^2}$ is t independent, then (2) means that g minimizes the "Action"

$$A(g) = A_{D,t_0,t_1}(g) = \frac{1}{2} \int_{t_0}^{t_1} \int_D |\partial_t g(t,x)|^2 dx dt$$

among all smooth trajectories $t \in [t_0, t_1] \to g$ on $G(D)$ satisfying

$$g(t_0) = g(t_0), \quad g(t_1) = g(t_1). \tag{3}$$

It turns out that the Euler equation are governed by the corresponding Least Action Principle on the configuration space $G(D)$, which means that Euler flows are just geodesic curves on this configuration space for the formal L^2 Riemannian structure. This has been known for a long time, in particular after a famous paper by Arnold [Ar]. (Among related results, let us quote [EM], [Sh], [She], [Se]....) Let us give a confirmation of the Least Action Principle through the following rigorous result.

Theorem 1.1 *Assume D to be the closure of a bounded convex open subset of the Euclidean space. Let (g,p) be a sufficiently smooth solution to the Euler equations (1) and let $[t_0,t_1]$ be a time interval short enough so that*

$$(t_1-t_0)^2\sum_{i,j=1}^{d}\partial^2_{x_ix_j}p(t,x)y_iy_j<\pi^2|y|^2$$

holds true for for all $t\in[t_0,t_1]$, $x\in D$, $y\neq 0$ in $\mathbf{R}^d$, then, for all smooth curve $t\to\tilde g(t)$ on the configuration space $G(D)$ such that

$$\tilde g(t_0)=g(t_0),\quad \tilde g(t_1)=g(t_1),$$

we have

$$A(g)\leq A(\tilde g),$$

where

$$A(g)=\frac{1}{2}\int_{t_0}^{t_1}\int_D|\partial_t g(t,x)|^2dxdt,$$

with equality if and only if g and $\tilde g$ coincide on $[t_0,t_1]$.

The proof is elementary and relies on the one-dimensional Poincaré inequality (that can be proved, using Fourier series, as an exercise)

Lemma 1.2 *Let $t_0<t_1$. For all absolutely continuous curves*

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

such that $z(t_0)=z(t_1)=0$ and z' is square integrable,

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

Proof

Let us compare g and $\tilde g$ subject to (3), fix $x\in D$ and denote $z(t)=g(t,x)$, $\zeta(t)=\tilde g(t,x)$. Since p is smooth, there is a constant $K=K(p)\geq 0$ such that

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

Since D is convex, the value of $K(p)$ can be taken as

$$\sup_{|y|=1}\sum_{i,j=1}^{d}\partial^2_{x_ix_j}p(t,x)y_iy_j.$$

By using the one dimensional Poincaré inequality, we get

$$\int_{t_0}^{t_1}|\zeta(t)-z(t)|^2dt \le \frac{(t_1-t_0)^2}{\pi^2}\int_{t_0}^{t_1}|\zeta'(t)-z'(t)|^2dt,$$

since $\zeta(t_j)=z(t_j)$ for $j=0,1$. Thus

$$\int_{t_0}^{t_1}[p(t,\zeta(t))-p(t,z(t))-\nabla p(t,z(t)).(\zeta(t)-z(t))]dt \le \int_{t_0}^{t_1}\frac{1}{2}|\zeta'(t)-z'(t)|^2dt,$$

provided that t_1-t_0 is small enough so that

$$\frac{(t_1-t_0)^2}{\pi^2}K(p)\le 1. \tag{4}$$

Since g is a solution to the Euler equations, we have

$$z''(t)=\partial^2_{tt}g(t,x)=-\nabla p(t,z(t)).$$

It follows, after integrating by part, that

$$\int_{t_0}^{t_1}[p(t,\zeta(t))-p(t,z(t))-z'(t).(\zeta'(t)-z'(t))]dt \le \int_{t_0}^{t_1}\frac{1}{2}|\zeta'(t)-z'(t)|^2dt,$$

which leads to

$$\int_{t_0}^{t_1}[-p(t,z(t))+\frac{1}{2}|z'(t)|^2]dt \le \int_{t_0}^{t_1}[-p(t,\zeta(t))+\frac{1}{2}|\zeta'(t)|^2]dt.$$

After integrating over $x\in D$, we get

$$\int_{t_0}^{t_1}\int_D[-p(t,g(t,x))+\frac{1}{2}|\partial_t g(t,x)|^2]dxdt$$
$$\le \int_{t_0}^{t_1}\int_D[-p(t,\tilde g(t,x))+\frac{1}{2}|\partial_t \tilde g(t,x)|^2]dxdt.$$

Since both g and $\tilde g$ are volume preserving

$$\int_{t_0}^{t_1}\int_D p(t,g(t,x))dxdt=\int_{t_0}^{t_1}\int_D p(t,\tilde g(t,x))dxdt=\int_{t_0}^{t_1}\int_D p(t,x)dxdt,$$

which shows that

$$\int_{t_0}^{t_1}\int_D\frac{1}{2}|\partial_t g(t,x)|^2dxdt \le \int_{t_0}^{t_1}\int_D\frac{1}{2}|\partial_t \tilde g(t,x)|^2dxdt$$

which completes the proof (the equality case being left as an exercise).

Remark

The Least Action Principle is satisfied only on sufficiently short time intervals. On larger time intervals, g is no longer a minimizer but rather a critical point of the Action. When D is not convex, the proof is valid with a larger constant $K(p)$ depending on the geometry of D. Condition (4) is sharp in the following case : D is the unique disk in $\mathbf{R}^2$, $t_0 = 0$, $t_1 = \pi$, $v(x) = (-x_2, x_1)$, $p(x) = \frac{1}{2}(x_1^2 + x_2^2)$ and $g(t,x) = xe^{it}$ (where complex notation $x = x_1 + ix_2$ is used). This fairly trivial solution to the Euler equations fails in minimizing the Action as soon as $t_1 > \pi$ (that is, after half a rotation of the disk).

1.4 The Shortest Path Problem

From a geometrical point of view (different from the natural PDE point of view which consists in addressing the Euler equations as an evolution equation with prescribed initial velocity field), it is natural to solve the Shortest Path Problem (SPP), or Least Action Problem (LAP). This problem amounts to minimize Action (2) among all smooth trajectories on $G(D)$ connecting two given elements g_0, g_1. Because of the group property of $G(D)$, we can assume g_0 to be the identity map I and denote g_1 by h. We can also rescale the time interval $[t_0, t_1]$ and set $t_0 = 0$, $t_1 = 1$.

Definition 1.3 *Given $h \in G(D)$, the Shortest Path Problem (SPP) consists in looking for a curve $t \in [0,1] \to g(t) \in G(D)$, such that $g(0) = I$, $g(1) = h$, minimizing the Action*

$$A(g) = A_D(g) = \frac{1}{2}\int_0^1 \int_D |\partial_t g(t,x)|^2 dx dt.$$

As mentioned before, the corresponding system of PDEs formally are the Euler equations, written in the so-called Lagrangian form

$$\partial_{tt}^2 g(t,x) = -\nabla p(t, g(t,x)),$$

with two point boundary conditions in time, which is different from solving the Cauchy problem, where only initial conditions are prescribed, namely $g(t=0,x)$ and $\partial_t g(t=0,x)$ for all $x \in D$.

1.5 A model for the SPP : The Closest Point Problem

1.5.1 Approximate geodesics

Before addressing directly the SPP, we can introduce a concept of approximate geodesics. The simplest way to define approximate geodesics on $G = G(D)$ is to introduce a penalty parameter $\epsilon > 0$ and to consider the formal dynamical system in the Hilbert space $H = L^2(D, \mathbf{R}^d)$

$$M'' + \frac{\delta}{\delta M}\left(\frac{d_H^2(M,G)}{2\epsilon^2}\right) = 0, \tag{5}$$

where the unknown M is a time dependent map in H, $\frac{\delta}{\delta M}$ denotes the gradient operator in H, and

$$d_H(M,G) = \inf_{g\in G} ||M-g||_H = \inf_{g\in S} ||M-g||_H = d_H(M,S), \tag{6}$$

is the distance in H between M and G, or, equivalently, between M and S, the L^2 closure of G, where $||.||_H$ is the Hilbert norm of H. This approach is similar-but not identical- to Ebin's slightly compressible flow theory [Eb], and is a natural extension of the theory of constrained finite dimensional mechanical systems [RU]. As the penalty parameter ϵ goes to zero, we expect that for appropriate initial data, typically for $M(t=0) = M_0 \in G$ and $M'(t=0) = v_0 \circ M_0$, where v_0 is a smooth divergence free vector field on D tangent to the boundary, the time dependent map M converges to a geodesic curve on G. Whenever a map M has a unique closest point $\pi_S(M)$ on S (which is not necessarily true since S is nor convex nor compact in H), the gradient of the squared distance from M to S is well defined at M and equal to $M - \pi_S(M)$. Thus, we may write, at least formally, the approximate geodesic equation (5) in the new form

$$\epsilon^2 M'' + M - \pi_S(M) = 0. \tag{7}$$

A rigorous analysis of this equation, introduced in [Br7], clearly requires an analysis of the Closest Point Problem (CPP) :

Definition 1.4 *The Closest Point Problem (CPP) amounts to find, given a map $M \in H = L^2(D, \mathbf{R}^d)$, a closest point (with respect to the L^2 norm) on the semi-group $S(D)$ of all measure preserving maps of D.*

The CPP can also be considered as a model problem for the SPP and will be subsequently investigated in details.

1.5.2 The time-discrete SPP

A different approach to the SPP is to use time discretization. Let $N > 2$ be a given integer. We call a semi-discrete shortest path a sequence $g_1, ..., g_N$ in $G(D)$, or more generally in the semi-group $S = S(D)$, that minimizes

$$A_N(g_1, ..., g_N) = \frac{1}{2} \sum_{i=2}^{N} ||g_i - g_{i-1}||^2_{L^2},$$

subject to the constraint

$$g_1 = I, \quad g_N = h,$$

where I denotes the identity map and h is the final configuration to be reached. A necessary optimality condition for such a sequence $g_1, ..., g_N$ is that for all $1 < i < N$, g_i minimizes

$$||g_{i-1} - g_i||^2_{L^2} + ||g_i - g_{i+1}||^2_{L^2},$$

or, equivalently (using that S is included in a sphere, which can be checked as an exercise),

$$||g_i - \frac{1}{2}(g_{i-1} + g_{i+1})||^2_{L^2}.$$

In other words, g_i is the closest point on S of the mid-point $\frac{1}{2}(g_{i-1} + g_{i+1})$. So, we see that, once again, the CPP naturally comes up. As a matter of fact, the CPP was originally introduced in [Br4] as the building block of a related numerical method to solve the Euler equations.

1.6 Non existence of solutions for the SPP

A local existence and uniqueness theorem for the SPP can be found in Ebin and Marsden paper [EM] : if h and I are sufficiently close in a sufficiently high order Sobolev norm, then there is a unique shortest path. In the large, uniqueness can fail for the SPP. For example, in the case when D is the unit disk, $g_0(x) = x = -g_1(x)$, the SPP has two solutions $g(t, x) = xe^{+i\pi t}$ and $g(t, x) = xe^{-i\pi t}$, where complex notations are used.

In 1985, A. Shnirelman [Sh] found, in the case $D = [0, 1]^3$, a class of data, that we will call "Shnirelman's class", for which the global SPP cannot have a (classical) solution. These data h are of the form

$$h(x_1, x_2, x_3) = (H(x_1, x_2), x_3),$$

where H is an area preserving mapping of the unit square, i.e. an element of $G([0,1]^2)$, such that

$$\delta_{[0,1]^3}(I,h) < \delta_{[0,1]^2}(I,H) < +\infty$$

(which means that the Action can be reduced if the third dimension motion is used). Indeed, let us consider a smooth curve g connecting I and h on $G([0,1]^3)$, generated by some smooth time-dependent divergence-free vector field $u(t,x)$, parallel to the boundary of D. Then, Shnirelman shows that there is such a curve $\tilde{g}$ satisfying

$$A_{[0,1]^3}(\tilde{g}) < A_{[0,1]^3}(g).$$

The new trajectory $\tilde{g}$ can be roughly obtained in two steps. First, u is rescaled by squeezing its vertical component (with symmetry with respect to $x_3 = 1/2$)

$$\tilde{u}_i(t,x) = u_i(t,x_1,x_2,2x_3), \quad i=1,2, \quad \tilde{u}_3(t,x) = \frac{1}{2}u_3(t,x_1,x_2,2x_3),$$

for $0 < x_3 < 1/2$, and

$$\tilde{u}_i(t,x) = u_i(t,x_1,x_2,2-2x_3), \quad i=1,2, \quad \tilde{u}_3(t,x) = \frac{1}{2}u_3(t,x_1,x_2,2-2x_3),$$

for $1/2 < x_3 < 1$. Next, the new field $\tilde{u}$, which is divergence-free and parallel to the boundary, but only Lipschitz continuous, is mollified and generates $\tilde{g}$. Of course, the vertical rescaling can be repeated *ad infinitum* in order to reduce the Action. This will generate infinitesimally small scales in the vertical direction.

So we can already guess that a good concept of generalized solutions to the SPP, for such data, must be related to the limit of the Euler equations under vertical rescaling, namely the so-called hydrostatic limit of the Euler equations discussed in [Li] (ch. 4.6).

1.7 The hydrostatic limit of the Euler equation

Let us consider the Euler equations in a thin domain such as $D = D_\epsilon = (\mathbb{R}^2/\mathbb{Z}^2) \times [0,\epsilon]$. For notational convenience, we denote the space variable by

$$(x,z) = (x_1,x_2,z) \in (\mathbb{R}^2/\mathbb{Z}^2) \times [0,\epsilon]$$

and the velocity by

$$(u, w) = (u_1, v_2, w) \in \mathbb{R}^3.$$

The vertical rescaling

$$u(t, x, z) \to u(t, x, z/\epsilon), \quad w(t, x, z) \to \epsilon w(t, x, z/\epsilon), \tag{8}$$

$$p(t, x, z) \to p(t, x, z/\epsilon),$$

leads, as $\epsilon \to$ to the so-called hydrostatic limit of the Euler equation [Li]

$$\partial_t u + (u.\nabla_x)u + w\partial_z u + \nabla_x p = 0. \tag{9}$$

$$\nabla_x.u + \partial_z w = 0, \quad \partial_z p = 0. \tag{10}$$

Although these equations look simpler than the original 3D Euler equations, they actually lead to considerable analytical difficulties, as mentioned in [Li]. To the best of our knowledge, only results in one horizontal space variable have been obtained for the initial value problem, under restrictive conditions on the initial conditions, related to the famous Rayleigh stability condition. Typically, it is required that the horizontal component u_1 of the velocity field at time $t = 0$ satisfies the local Rayleigh condition

$$\partial^2_{zz} u_1(t = 0, x_1, z) > 0$$

for all (x_1, z). Under additional conditions, existence and uniqueness of local smooth solutions were proved by the author in [Br6]. Convergence from the original Euler equations was proved by Grenier in [Gr] under similar conditions. In addition, Grenier showed that some solutions of the Euler equations may not converge to the hydrostatic solutions if they do not satisfy the local Rayleigh condition at $t = 0$. A key step of the analysis provided in [Br6] is an appropriate reformulation of the hydrostatic equations. Let us consider a smooth solution (u, w, p) of the hydrostatic equations and define a Lagrangian foliation to be a family of sheets $z = Z(t, x, a)$, labeled by $a \in [0, 1]$, where Z is a smooth function such that :

$$0 \le Z(t, x, a) \le 1, \quad Z(t, x, 0) = 0, \quad Z(t, x, 1) = 1, \tag{11}$$

$$\partial_a Z(t, x, a) > 0, \tag{12}$$

$$\partial_t Z(t, x, a) + u(t, x, Z(t, x, a)).\nabla_x Z(t, x, a) = w(t, x, Z(t, x, a)). \tag{13}$$

If $Z(0,x,a)$ is given and compatible with (11), (12), it is always possible to get a Lagrangian foliation at least on a short interval of time. Then, the hydrostatic equations become (after elementary computations)

$$\partial_t c + \nabla_x.(cv) = 0, \tag{14}$$

$$\partial_t(cv) + \nabla_x.(cv \otimes v) + c\nabla_x p = 0, \tag{15}$$

$$\int_0^1 c(t,x,a)da = 1, \tag{16}$$

where the new unknowns (c,v) are defined by :

$$c(t,x,a) = \partial_a Z(t,x,a) > 0, \tag{17}$$

$$v(t,x,a) = u(t,x,Z(t,x,a)) \tag{18}$$

and p is unchanged. This change of variable turns out to be essentially the one we need to solve the SPP in a generalized sense.

1.8 The Relaxed Shortest Path Problem

To solve, in a generalized sense, the SPP, in particular for data h in Shnirelman's class, a natural idea is to introduce appropriate "Young's measures" [Yo], [Ta]. There are different ways [Br1], [She], [Sh2]. One approach can be used, as in [Br2], which turns out to be closely related to the hydrostatic rescaling of the Euler equations. Given a smooth trajectory $t \in [0,1] \to g(t)$ on $G(D)$, we define two measures (respectively nonnegative and vector-valued)

$$c(t,x,a) = \delta(x - g(t,a)), \quad m(t,x,a) = \partial_t g(t,a)\delta(x - g(t,a)), \tag{19}$$

defined on $Q' = [0,1] \times D \times D$. These measures satisfy

$$\int_A c(t,x,da) = 1, \tag{20}$$

$$\partial_t c + \nabla_x.m = 0, \tag{21}$$

$$c(0,x,a) = \delta(x-a); \quad c(T,x,a) = \delta(x - h(a)). \tag{22}$$

Moreover, m is absolutely continuous with respect to c, with a vector-valued density $v \in L^2(Q', dc)^d$, so that $m = cv$, and the Action is given by

$$A(g) = \int_{D\times A} \frac{1}{2}|v(t,x,a)|^2 c(t,dx,da), \tag{23}$$

or, equivalently, $A(g) = K(c, m)$ where

$$K(c, m) = \sup_{(F,\Phi)} \int_{Q'} (F(t, x, a)dc(t, x, a) + \Phi(t, x, a).dm(t, x, a)), \tag{24}$$

and the supremum is taken among all continuous functions F and Φ on Q', with values respectively in $\mathbb{R}$ and $\mathbb{R}^d$, such that

$$F(t, x, a) + \frac{1}{2}|\Phi(t, x, a)|^2 \leq 0, \tag{25}$$

pointwise. Then a natural definition of the *Relaxed* SPP, called RSPP, is to look for pairs of measures (c, m) that minimize $K(c, m)$ and are admissible in the sense of (20),(21) and (22), but do not necessarily satisfy (19).

1.9 Solutions of the Relaxed Shortest Path Problem

In [Br2], it is shown that, for $D = [0, 1]^d$ and each data $h \in S(D)$ (which of course includes Shnirelman's class), the RSPP always has solutions (c, m) and that there exists a unique locally bounded measure $\nabla_x p(t, x)$ in the interior of $Q = [0, 1] \times D$, depending only on h, such that

$$\partial_t(cv) + \nabla_x.(cv \otimes v) + \underline{c}\nabla_x p = 0, \tag{26}$$

holds in the sense of distributions on the interior of Q'. In this equation, $\underline{c}$ is an appropriate extension of c, allowing the (non-obvious) pairing with $\nabla_x p$. More precisely, for any fixed $e \in \mathbb{R}^d$ and any nonnegative smooth radial compactly supported mollifier γ on $\mathbb{R}^d$,

$$\underline{c}(t, x, a) = \lim_{\delta \to 0} \int_{-1/2}^{+1/2} d\theta \int c(t, x - 2\theta\delta e - \delta y, a)\gamma(y)dy, \tag{27}$$

for the weak-* topology of the dual space of $L^1(|\nabla p|, C(D))$ (the vector space of all $|\nabla p|$ integrable function of $(t, x) \in Q$ with values in the space $C(D)$ of all continuous function on D). So, we have exactly recovered the hydrostatic limit of the Euler equations, in their second formulation, namely (14), (15), (16), as the optimality condition of the RSPP.

This result is obtained in several steps. First, the existence part is easily obtained through standard convex analysis and duality theory. Next, a priori bounds are obtained for the Lagrange multiplier of constraint (20), namely

$\nabla_x p$, which turns out to be uniquely defined by a duality argument. Finally the conservation of momentum (26) is obtained as an optimality condition.

Still in [Br2], the original SPP and the RSPP are related in the case $D = [0,1]^3$. It is shown for *any* data $h \in S([0,1]^3)$ of the form

$$h(x_1, x_2, x_3) = (H(x_1, x_2), x_3),$$

and, in particular, for any data in Shnirelman's class, that, for any $\epsilon > 0$, there is a smooth trajectory $t \in [0,1] \to g_\epsilon(t)$ on $G(D)$ such that

$$g_\epsilon(0) = I, \quad A(g_\epsilon) + ||g_\epsilon(1,.) - h||^2_{L^2(D)} \leq I(h) + \epsilon. \tag{28}$$

where $I(h)$ is the optimal value of the RSPP. In addition, the measures (c_ϵ, m_ϵ) associated with g_ϵ, through (19), converge, as $\epsilon \to 0$ to the generalized solutions of the RSPP. Moreover, the g_ϵ are almost solution of the Euler equations in the sense that their velocity field v_ϵ satisfy

$$\nabla.v_\epsilon = 0, \quad \partial_t v_\epsilon + (v_\epsilon.\nabla)v_\epsilon \to -\nabla p,$$

weakly, as ϵ tends to zero.

1.10 Consistency of the relaxed SPP with the Euler equations.

A local consistency result of the relaxed SPP with the classical Euler equations is also provided in [Br2] :

Theorem 1.5 *Let (g,p) be a smooth solution to the Euler equations $T > 0$ such that $\Lambda T^2 < \pi^2$, where Λ is the supremum on Q of the largest eigenvalue of the hessian matrix of p, and set $h = g(T)$. Then, the pair (c,m) associated with g through (19) is the unique solution of the relaxed SPP.*

1.11 Example of generalized solutions

Explicit examples of non trivial generalized solutions to the weak SPP were first described in [Br1]. Let us just quote a typical example, when D is the unique disk and $h(x) = -x$. Then, the classical SPP has two distinct solutions $g_+(t,x) = e^{i\pi t}x$ and $g_-(t,x) = e^{-i\pi t}x$, with the same pressure field

$p = \pi^2|x|^2/2$, where complex notations are used. A generalized solution (c, m) is given by

$$\int_{Q'} f(t,x,a)dc(t,x,a) = \int_{[0,1]\times D}\int_0^1 f(t, G(t,a,\theta), a)d\theta dt da, \qquad (29)$$

$$\int_{Q'} f(t,x,a)dm(t,x,a) = \int_{[0,1]\times D}\int_0^1 \partial_t G(t,a,\theta) f(t, G(t,a,\theta), a)d\theta dt da, \qquad (30)$$

for all continuous function f, where

$$G(t,a,\theta) = a\cos(\pi t) + (1-|a|^2)^{1/2} e^{2i\pi\theta}\sin(\pi t) \in D.$$

This generalized solution describes a very peculiar wave-like motion of the fluid particles. Each particle initially located at $a \in D$ splits up along a circle of radius $(1-|a|^2)^{1/2}\sin(\pi t)$, with center $a\cos(\pi t)$, that moves across the unit disk and shrinks down to the point $-a$ when $t = 1$. Of course, its acceleration is still given by the pressure field, $p = \pi^2|x|^2/2$, as expected from the theory of the RSPP.

2 Measure preserving maps and density theorems

In this second section, measure preserving maps are studied in a relatively general framework. Density theorems are also discussed.

2.1 General definition

Definition 2.1 *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 $\beta-$ integrable on Y, $x \to f(\phi(x))$ is Borel and $\alpha-$ integrable on X and

$$\int_X f(\phi(x))d\alpha(x) = \int_Y f(y)d\beta(y). \tag{31}$$

2) Of course, the definition can be extended to abstract measure spaces.
3) In the case $X = Y = [0,1]$, $\alpha = \beta = |.|$, where $|.|$ denotes the Lebesgue measure, some examples of measure preserving maps are given by

$$\phi(x) = x \tag{32}$$

$$\phi(x) = x + \frac{1}{2}, \quad mod.1 \tag{33}$$

(which is discontinuous),

$$\phi(x) = 1 - x \tag{34}$$

(which is orientation reversing),

$$\phi(x) = \min(2x, 2 - 2x) \tag{35}$$

(which is not one-to-one).

4) A remarkable theorem (see [Ro] 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 (31). 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 $([0,1], |.|)$ as long as X is metric, separable, complete without any point of positive α measure.

Exercise

Discuss the case of two discrete space X, Y, with respectively N and M elements. In the particular case when $X = Y$ and $\alpha = \beta$ is the counting measure, describe the set of all measure preserving maps.

2.2 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 $|.|$ 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 (35) 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 Lebesgue measurable functions are always almost everywhere equal to a Borel functions, so that there is no problem to define S as a subset of L^p.)

Exercise

Show 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 1, because of the general identity

$$\partial_t \log \det(\partial_x g_t(v)(x)) = (\nabla.v)(t, g_t(v)(x)), \tag{36}$$

valid for all smooth vector fields v. In particular, $g_t(v)$ is (Lebesgue) measure preserving, because of the change of variable formula

$$\int_D f(\phi(x)) \det(\partial_x g_t(v)(x)) dx = \int_D f(x) dx, \quad \forall f.$$

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 (36) using that $\det(I + A) = 1 + tr(A) + O(A^2)$.

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 [EG] is recommanded.) Can such map be smooth (at least C^1) without being one-to-one?

2.3 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 2.2 *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) As D has finite measure, all L^p topologies ($p < \infty$) are equivalent.
2) Clearly $d \geq 2$ is needed !
3) 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 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 [AK]...) 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.

2.4 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 very much from the one provided in Neretin's paper [Ne].

2.4.1 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. This suggests 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

$$\phi(x) = x - x_{n,i} + x_{n,\sigma(i)}, \tag{37}$$

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 2.3 *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.*

2.4.2 Proof of Proposition 2.3

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 2.4 *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 $\{s_1 \quad s_2 \; ; \; s_1 \in S_1 \; , \quad s_2 \in S_2\}$.*

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 2.3, it is now enough to approximate by a sequence in G_0 a permutation of two adjacent subcubes. 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

$$(t,x) \in]0,1[\times Q \to v(t,x) \in \mathbf{R}^d$$

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 [Ne].) Notice that $d \geq 2$ is clearly needed to achieve the construction.

2.4.3 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 prova 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 2.5 *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 margins are respectively α and β, namely*

$$\mu(A \times Y) = \alpha(A), \quad \mu(X \times B) = \beta(B),$$

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

$$\mu(A \times B) = \alpha(\phi^{-1}(B) \quad A),$$

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.

2.4.4 Density of P in S and DS

Let us now return to the case $X = Y = D = [0,1]^d$ and $\alpha = \beta = |.|$. 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 2.6 *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 from which 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. This, 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

$$|I_1 - I_2| \leq \sup |f| 2^{(2n-l)d} + 3\eta(2^{-n-l+d/2}),$$

car $L = 2^{ld}$, $K = N^2 = 2^{2nd}$. This completes the proof, after letting first l and then n to $+\infty$.

2.4.5 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.

2.5 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 kind of Lusin theorem is also known for MPM that ϕ are one-to-one in the almost everywhere sense. 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).

3 The Closest Point Problem and the Polar Factorization of Maps

Since $S = S(D)$ is not convex (because S is included in a sphere), the CPP is not trivial. (Notice that the CPP on $G(D)$ is even worse since this subset of $S(D)$ is not closed). However, since S is a closed bounded subset of a Hilbert space, it follows from Edelstein's theorem [Au] that almost every element $M \in L^2(D, \mathbf{R}^d)$, in the topological sense of Baire category theorem, has a unique closest point on $S(D)$. However, such a result is quite abstract and one of our first task will be to address this problem more concretely.

This will lead to a Polar Factorization theorem for maps in the Hilbert space $L^2(D, \mathbf{R}^d)$, involving the semi-group $S(D)$ (rather than $G(D)$) and the "dual" convex cone

$$K(D) = \{M \in L^2(D, \mathbf{R}^d) \; ; \quad ((M, I - h)) \geq 0, \quad \forall h \in S(D)\},$$

where $((.,.))$ denotes the L^2 inner product. This convex cone will be characterized as the set of all square integrable mappings from D into $\mathbf{R}^d$ that coincide almost everywhere on D with the subgradient of some semi lower continuous convex function defined on $\mathbf{R}^d$. More precisely, we will show, following [Br3],

Theorem 3.1 *Assume that $M \in L^2(D, \mathbf{R}^d)$ satisfies the following non degeneracy condition : if N is a negligible subset of D, then $M^{-1}(N)$ is also negligible. Then there is a unique decomposition*

$$M(x) = \nabla\Phi(h(x)), \quad a.e. \; x \in D,$$

where h belongs to $S(D)$ and Φ (defined up to an additive constant) is the restriction to D of a lower semi continuous convex function on $\mathbf{R}^d$. Moreover, h is the unique closest point to M on D, $\nabla\Phi$ is the unique rearrangement of M in the class $K(D)$.

By rearrangement of M, we mean any map v from D into $\mathbf{R}^d$ such that

$$\int_D \phi(v(x))dx = \int_D \phi(M(x))dx$$

holds for all $\phi \in C_c(\mathbf{R}^d)$. Theorem 3.1 shows that a map $M \in L^2$ has a unique rearrangement as a gradient of some convex potential, which generalizes the classical theory on non decreasing rearrangements of real valued functions [HLP]. Theorem 3.1 can also be seen as a non linear Hodge decomposition theorem. Indeed, when *formally* linearized about the identity map, the polar decomposition yields the classical unique decomposition of vector fields

$$z = w + \nabla p,$$

where z is a given vector field, w a divergence free vector field, parallel to the boundary of D, and p is a real valued function.

Exercise

Obtain the Hodge decomposition by a formal linearization of the polar factorization.

Exercise

Find the polar factorization of a linear map when D is a ball. (Assume the uniqueness of the factorization and use the classical polar factorization theorem for real square matrices.) What happens if D is not a ball?

Regularity results were obtained by Caffarelli [Ca1] [Ca2].

Theorem 3.2 *Assume M to have $C^{1,\alpha}$ regularity on D up to the boundary, for some $0 < \alpha < 1$, D and $M(D)$ to be strictly convex with a smooth boundary, and the jacobian determinant of M to be positive and bounded away from zero (which insures the non degeneracy condition). Then Φ is strictly convex on D and both $\nabla\Phi$ and h belongs to $C^{1,\alpha}$, up to the boundary. Moreover h belongs to $G(D)$ and Φ can be recovered by solving a Monge-Ampère equation.*

This regularity result shows that, under strong assumptions on D and M, M has a unique L^2 closest point on $G(D)$. The proof is based on the fact that the Legendre-Fenchel transform of Φ, namely

$$\Psi(y) = \sup_{x \in \mathbf{R}^d} (x.y - \Phi(x)),$$

is a weak solution to the Monge-Ampère equation

$$\det D^2\Psi = \rho,$$

where $\rho(x)dx$ is the image measure of dx by M. Caffarelli shows that Ψ is a solution in the sense of Alexandrov and is strictly convex. Then, he obtains both local [Ca1] and global [Ca2] regularity.

3.1 The Monge-Kantorovich theory

The proof of Theorem 3.1 relies on a relaxation technique called Monge-Kantorovich theory. Its origin goes back to Monge's mass transfer problem addressed in the 'mémoire sur la théorie des déblais et des remblais' [Mo]. A modern approach, based on duality arguments, is due to Kantorovich [Ra] and was used in Probability theory by Vershik, Sudakov, and more recently Rachev [Ra, Su, Ve]. Let us quote a typical result (which does not differ essentially from Theorem 3.1).

Theorem 3.3 *Assume ρ_0 and ρ_1 to be two nonnegative Lebesgue integrable compactly supported functions on $\mathbf{R}^d$, such that*

$$\int_{\mathbf{R}^d} \rho_0(x)dx = \int_{\mathbf{R}^d} \rho_1(x)dx = 1.$$

Then there is a Lipschitz continuous convex function Φ on $\mathbf{R}^d$ such that

$$\int_{\mathbf{R}^d} f(\nabla\Phi(x))\rho_0(x)dx = \int_{\mathbf{R}^d} f(x)\rho_1(x)dx$$

holds for any continuous function f on $\mathbf{R}^d$.

Sketch of the proof

Let us consider a ball B in $\mathbf{R}^d$ containing the supports of both ρ_0 and ρ_1 and introduce the set M of all Borel regular probability measures ν on $B \times B$ having $\rho_0(x)dx$ and $\rho_1(x)dx$ as margins, which means

$$\int_{B\times B} f(x)\nu(dx,dy) = \int_B f(x)\rho_0(x)dx,$$
$$\int_{B\times B} f(y)\nu(dx,dy) = \int_B f(y)\rho_1(y)dy,$$

for all continuous functions f on $\mathbf{R}^d$. By using Riesz representation theorem on Borel measures and elementary convex analysis (as the Rockafellar theorem stated in [Brez]), we obtain the duality equality

$$\max_{\nu\in M} \int_{B\times B} x.y \ \nu(dx,dy) = \inf \int_B [\Phi(x)\rho_0(x) + \Psi(x)\rho_1(x)]dx,$$

where the infimum is taken over all pairs (Φ, Ψ) of continuous functions on B satisfying

$$\Phi(x) + \Psi(y) \geq x.y, \quad \forall x \in B, \quad \forall y \in B.$$

Then, it can be established that the infimum is attained by a pair (Φ, Ψ) such that Φ is the restriction of a Lipschitz continuous convex function defined on $\mathbf{R}^d$, and for $\rho_0(x)dx$ almost every point of $\mathbf{R}^d$, Ψ coincide with the Legendre-Fenchel transform of Φ,

$$LF(\Phi)(y) = \sup_{x \in \mathbf{R}^d} (x.y - \Phi(x)).$$

Moreover, if $\nu = \nu_{opt} \in M$ maximizes $\int_{B \times B} x.y \ \nu(dx, dy)$, then

$$\Phi(x) + \Psi(y) = x.y$$

holds for $\nu_{opt}-$ almost every $(x, y) \in \mathbf{R}^d \times \mathbf{R}^d$. Using well known properties of the Legendre-Fenchel transform, one deduces that ν_{opt} is necessarily of the form

$$\nu_{opt}(dx, dy) = \delta(y - \nabla\Phi(x))\rho_0(x)dx$$

which implies

$$\int_{\mathbf{R}^d \times \mathbf{R}^d} f(y)\nu_{opt}(dx, dy) = \int_{\mathbf{R}^d} f(\nabla\Phi(x))\rho_0(x)dx,$$

for all continuous function f on $\mathbf{R}^d$ and achieves the proof since the second margin of ν_{opt} is $\rho_1(x)dx$.

Remark 1

We can define the Kantorovich-Wasserstein distance (see [RR] for example) between ρ_0 and ρ_1 by setting

$$\Delta(\rho_0, \rho_1) = \inf_{\nu \in M} (\int_{D \times D} |x - y|^2 \nu(dx, dy))^{1/2}. \quad (38)$$

Then we get

$$\int_D |\nabla\Phi(x) - x|^2 \rho_0(x)dx = \Delta(\rho_0, \rho_1)^2.$$

Indeed,

$$\int_D |\nabla\Phi(x) - x|^2 \rho_0(x)dx = \int_{D \times D} |y - x|^2 \nu_{opt}(dx, dy)$$

$$= \int_D |x|^2(\rho_0(x) + \rho_1(x))dx - \int_{D\times D} 2y.x\ \nu_{opt}(dx, dy)$$

(since ρ_0 and ρ_1 are the margins of ν_{opt})

$$\leq \int_D |x|^2(\rho_0(x) + \rho_1(x))dx - \int_{D\times D} 2y.x\ \nu(dx, dy)$$

for every $\nu \in M$ (since ν_{opt} maximizes $\int y.x\ \nu(dx, dy)$),

$$= \int_{D\times D} |y - x|^2 \nu(dx, dy)$$

(since ρ_0 and ρ_1 are also the margins of ν).

Remark 2

The proof of Theorem 3.1 uses similar arguments and corresponds to the special case where $\rho_0(x) = 1$ and $\rho_1(x)dx$ is the image measure of dx by the mapping u. However the proof is more complicated, partly due to the assumption that u belongs to L^2, which rules out the assumption that ρ_1 is compactly supported.

4 Main steps of the study of the relaxed SPP.

The detailed proofs can be found in [Br2]. It is first deduced from the Rockaffelar theorem in classical convex analysis [Brez]

Proposition 4.1 *The infimum of the RSPP $I^*(h)$ is always achieved and for every $\epsilon > 0$, there exist some continuous functions $\phi_\epsilon(t,x,a)$ on Q' and $p_\epsilon(t,x)$ on Q, with $\partial_t\phi_\epsilon$, $\nabla_x\phi_\epsilon$ continuous on Q' and $\int_D p_\epsilon(t,x)dx = 0$, such that, for every optimal solution (c,m),*

$$\partial_t\phi_\epsilon + \frac{1}{2}|\nabla_x\phi_\epsilon|^2 + p_\epsilon \leq 0 \tag{39}$$

and

$$\int_{Q'}(|\partial_t\phi_\epsilon + \frac{1}{2}|\nabla_x\phi_\epsilon|^2 + p_\epsilon| + |v - \nabla_x\phi_\epsilon|^2)dc \leq \epsilon^2. \tag{40}$$

Then, an approximate regularity result is obtained

Proposition 4.2 *Let $0 < \tau < T/2$ and $Q'_\tau = [\tau/2, T-\tau/2] \times D \times A$. Let $x \in D \to w(x) \in \mathbb{R}^d$ be a smooth divergence-free vector field, parallel to ∂D, and $s \in \mathbb{R} \to e^{sw}(x) \in D$ be the integral curve of w passing through x at $s = 0$. Then,*

$$\int_{Q'_\tau}|\nabla_x\phi_\epsilon(t,x,a) - v(t,x,a)|^2 dc(t,x,a) \leq C\epsilon^2, \tag{41}$$

$$\int_{Q'_\tau}|\nabla_x\phi_\epsilon(t,x,a)|^2 dc(t,x,a) \leq C, \tag{42}$$

$$\int_{Q'_\tau}|\nabla_x\phi_\epsilon(t+\eta, e^{\delta w}(x), a) - \nabla_x\phi_\epsilon(t,x,a)|^2 dc(t,x,a) \leq (\epsilon^2 + \eta^2 + \delta^2)C, \tag{43}$$

for all optimal solution (c,m) and all η, δ and $\epsilon > 0$ small enough, where C depends only on D, T, τ and w.

Since $c(t,x,a)$ is a measure, possibly highly concentrated (like a delta measure in x, as in the case of classical solutions), it is unclear how to deduce from Proposition 4.2 a bound such as

$$\int_{Q'_\tau}(|\partial_t v|^2 + |\nabla_x v|^2)dc \leq C, \tag{44}$$

by letting first $\epsilon \to$ (to get v instead of $\nabla_x \phi_\epsilon$), then $\delta, \eta \to 0$. Such a bound would be meaningful, if c could be bounded away from zero, which is exactly the contrary of the classical case and cannot be expected, anyway, because of the initial and final conditions. However a bound on $\int |\nabla p|$ is expectable. The formal (and, of course, not rigorous) argument is as follows. Starting from (40), we get

$$(\partial_t \phi + \frac{1}{2}|\nabla_x \phi|^2 + p)c = 0.$$

Differentiating in x, we formally get (26) or

$$(\partial_t v + (v.\nabla_x)v + \nabla_x p)c = 0.$$

Then, integrating in $a \in A$,

$$\int_A (\partial_t v + (v.\nabla_x)v)c(t,x,da) = -\nabla_x p,$$

and, by Schwarz inequality,

$$(\int |\nabla_x p|)^2 \leq \int |\partial_t v|^2 dc \int dc + \int |\nabla_x v|^2 dc \int |v|^2 dc.$$

All these calculations are incorrect. However, the formal idea can be made rigorous, by working only on the ϕ_ϵ and using finite differences instead of derivatives, and lead to

Theorem 4.3 *The family (∇p_ϵ) converges in the sense of distributions toward a unique limit ∇p, depending only on h, which is a locally bounded measure in the interior of Q and is uniquely defined by*

$$\nabla p(t,x) = -\partial_t \int v(t,x,a)c(t,x,da) - \nabla_x . \int (v \otimes v)(t,x,a)c(t,x,da), \quad (45)$$

for ALL optimal solution $(c, m = cv)$.

Finally, (26) is established.

4.1 Existence of admissible solutions.

In this subsection we obtain admissible solutions to the RSPP through an explicit construction closely related to the one introduced in [Br1] for generalized flows on the torus $\mathbb{T}^d$ (and used later in [Sh2]). We only perform the construction in the cases $D = \mathbb{T}^d$ and $D = [0,1]^d$, the second one being an extension of the first one.

Admissible solutions on the torus.

Let $h \in S(D)$ where $D = \mathbb{T}^d$. We introduce, for $(x,y,z) \in D^3$, a curve $t \in [0,T] \rightarrow \omega(t,x,y,z) \in D$, made, for $0 \leq t \leq T/2$, of a shortest path (with constant speed) going from x to y on the torus $\mathbb{T}^d$, and, for $T/2 \leq t \leq T$, of a shortest path (with constant speed) going from y to z. Such a curve is uniquely defined for Lebesgue almost every $(x,y,z) \in D^3$. Then, we set, for every continuous function f on $Q' = [0,T] \times D \times D$,

$$< c, f >= \int_{Q'} f(t, \omega(t,a,y,h(a)), a) dt dy da$$

$$< m, f >= \int_{Q'} \partial_t \omega(t,a,y,h(a)) f(t, \omega(t,a,y,h(a)), a) dt dy da.$$

(Intuitively, this amounts to define a generalized flow for which each particle a is first uniformly scattered on the torus at time $T/2$ and then focused to the target $h(a)$ at time T.) This makes (c,m) an admissible solution. Let us check, for instance, that the continuity equation is sqtisfied, by considering a continuous function f that does not depend on a and showing that $< c, f >= \int_Q f$. We split $< c, f >= I_1 + I_2$ according to $t \leq T/2$ or not. For $t \geq T/2$, we have, by definition, $\omega(t,a,y,h(a)) = \omega(t,y,y,h(a))$. Since we work on the torus $\mathbb{T}^d$, $\omega(t,y,y,h(a)) = y + \omega(t,0,0,h(a)-y)$. Thus

$$I_2 = \int_{[T/2,T]\times D\times D} f(t, y + \omega(t,0,0,x-y)) dx dy dt$$

(since $a \in D \rightarrow x = h(a) \in D$ is Lebesgue measure preserving)

$$= \int_{[T/2,T]\times D\times D} f(t, y + \omega(t,0,0,x)) dx dy dt = \int_{[T/2,T]\times D\times D} f(t,y) dx dy dt$$

$$= \int_{[T/2,T]\times D} f(t,y) dy dt$$

(by using twice the translation invariance of the Lebesgue measure on the torus) and, doing the same for I_1, we conclude that $< c, f >= \int_Q f$. We also get, for $K(c,m)$ the following estimate that does not depend on the choice of $h \in S(D)$

$$2K(c,m) = \int_{[0,T]\times D\times D} |\partial_t \omega(t,a,y,h(a))|^2 da dy dt$$

$$= T/2 \int_{D\times D} ((d_D(a,y)/(T/2))^2 + (d_D(y,h(a))/(T/2))^2) da dy$$

$$= \frac{4}{T} \int_{D\times D} d_D(x,y)^2 dx dy \le d/T$$

(where $d_D(.,.)$ denotes the geodesic distance on the torus).

Admissible solutions on the unit cube.

Let us now lift the unit cube to the torus by shrinking it by a factor 2 and reflecting it 2^d times through each face of its boundary. To do that, we introduce the Lipschitz continuous mapping

$$\Theta(x) = 2(\min(x_1, 1-x_1), ..., \min(x_d, 1-x_d)), \tag{46}$$

from $[0,1]^d$ onto $[0,1]^d$, and its 2^d reciprocal maps, each of them being denoted by Θ_k^{-1}, with $k \in \{0,1\}^d$, and mapping back $[0,1]^d$, one-to-one, to the cube $2^{-1}(k+[0,1]^d)$. Given $h \in S([0,1]^d)$, we associate $\tilde{h} \in S(\mathbb{T}^d)$ by setting

$$\tilde{h}(x) = \Theta_k^{-1}(h(\Theta(x)),$$

when $x \in \frac{1}{2}(k+[0,1]^d)$, $k \in \{0,1\}$. We consider the admissible pair $(\tilde{c}, \tilde{m})$ associated with $\tilde{h}$ and constructed exactly as in the previous subsection. Then we set

$$c(t,x,a) = \frac{1}{2^d} \sum_{k\in\{0,1\}^d} c(t, \Theta_k^{-1}(x), \Theta_k^{-1}(a)),$$

$$v(t,x,a) = \frac{1}{2^d} \sum_{k\in\{0,1\}^d} [((v(t,\Theta_k^{-1}(x), \Theta_k^{-1}(a))).\nabla_x)\Theta](\Theta_k^{-1}(x)).$$

Explicit calculations show that (c,m) is an admissible solution for h and that (as in [Sh2])

$$K(c,m) \le \frac{2d}{3T} \le dT^{-1}.$$

Notice that it is important to use a continuous transform Θ (so that the particle trajectories are properly reflected at the boundary of the unit cube and not broken) in order to keep the continuity equation, but there is no need to use a one-to-one transform, which makes the construction very easy.

References

[Ar] V.I.Arnold, *Ann. Institut Fourier 16 (1966) 319-361.*

[AK] V.I.Arnold, B.Khesin, *Topological methods in Hydrodynamics, Springer Verlag, 1998.*

[Au] J.-P. Aubin, *Mathematical Methods of Game and Economic Theory. North-Holland, Amsterdam, 1979.*

[BN] S. Baigent, J. Norbury, *Two discrete models for semi-geostrophic dynamics, Phys. D 109 (1997) 333-342.*

[Br1] Y.Brenier, *J.of the AMS 2 (1990) 225-255.*

[Br2] Y.Brenier, *Minimal geodesics on groups of volume-preserving maps, Comm. Pure Appl. Math. 52 (1999) 411-452.*

[Br3] Y.Brenier, *Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. 44 (1991) 375-417.*

[Br4] Y.Brenier, *Comp. Meth. in Appl. Mech. 75 (1989) 325-332.*

[Br5] Y.Brenier, *The Dual Least Action Problem for an Ideal, Incompressible fluid. Arch. Rational Mech. (1993).*

[Br6] Y.Brenier, *Homogeneous hydrostatic flows with convex velocity profiles, Nonlinearity 12 (1999), no. 3, 495–512. Arch. Rational Mech. (1993).*

[Br7] Y.Brenier, *Derivation of the Euler equations from a caricature of Coulomb interaction, Comm. Math. Phys. 212 (2000), no. 1, 93–104.*

[Brez] H.Brezis, *Analyse fonctionnelle.*

[Ca1] L.Caffarelli, *J.of the AMS 2 (1992) .*

[Ca2] L. Caffarelli, *Boundary regularity of maps with convex potentials, Comm. Pure Appl. Math. 45 (1992) 1141-1151.*

[Ch] J.-Y.Chemin, *Fluides parfaits incompressibles, Astérisque 230 (1995).*

[CNP] M. Cullen, J. Norbury, J. Purser, *Generalised Lagrangian solutions for atmospheric and oceanic flows, SIAM J. Appl. Math. 51 (1991), 20-31.*

[DM] R.DiPerna, A.Majda *Comm. Math. Phys. 108 (1987), 667-689.*

[Eb] D. Ebin, *The motion of slightly compressible fluids viewed as a motion with strong constraining force,* Ann. of Math. *(2) 105 (1977) 141-200.*

[EM] D.Ebin, J.Marsden, *Ann. of Math. 92 (1970) 102-163.*

[ER] Y.Eliashberg, T.Ratiu, *Invent. Math. 103 (1991) 327-340.*

[Eu] L.Euler, *Opera Omnia, Series Secunda, 12, 274-361.*

[EG] L.C. Evans, R. Gariepy, *Measure theory and fine properties of functions, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.*

[Gr] E. Grenier, *On the derivation of homogeneous hydrostatic equations,* it M2AN Math. Model. Numer. Anal. 33 (1999), no. 5, 965–970.

[HLP] G.H.Hardy, J.E.Littlewood, G.Polya, *Inequalities. Cambridge Univ. Press, 1952.*

[Ho] B. Hoskins, *The mathematical theory of frontogenesis, Annual review of fluid mechanics, Vol. 14, pp. 131-151, Palo Alto, 1982.*

[La] P. Lax, *Approximation of measure preserving transformations, Comm. Pure Appl. Math. 24 (1971) 133-135.*

[Lic] L.Lichtenstein, *Math. Z., volumes 23-32 (1925-1932).*

[Li] P.-L. Lions, *Mathematical topics in fluid mechanics. Vol. 1. Incompressible models, Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, New York, 1996.*

[Ma] A.Majda, *Proceedings of the International Congress of Mathematicians, Kyoto 1990, Springer, 1991.*

[MP] C.Marchioro, M.Pulvirenti, *Mathematical theory of incompressible nonviscous fluids, Springer, New York, 1994.*

[Mo] G.Monge, *Mém. Math. Phys. Acad. Roy. Sci. Paris (1781), 666-704.*

[Ne] Y. Neretin, *Categories of bistochastic measures and representations of some infinite-dimensional groups, Sb. 183 (1992), no. 2, 52-76.*

[Ra] S.T.Rachev, *Th. Prob. Appl. 49 (1985) 647-676.*

[Ro] H.L. Royden, *Real analysis, Macmillan Publishing Company, New York, 1988.*

[RU] H. Rubin, P. Ungar, *Motion under a strong constraining force, Comm. Pure Appl. Math. 10 (1957) 65-87.*

[RR] L.Rüschendorf, S.T.Rachev, *J. of Multivariate Analysis 32 (1990) 48-54.*

[Se] D. Serre, *Sur le principe variationnel des équations de la mécanique des fluides parfaits, RAIRO Model. Math. Anal. 27 (1993) 739-758.*

[She] V. Shelukhin, *Existence theorem in the variational problem for compressible inviscid fluids, Manuscripta Math. 61 (1988) 495-509.*

[Sh] 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.*

[Sh2] A. I. Shnirelman, *Generalized fluid flows, their approximation and applications, Geom. Funct. Anal. 4 (1994) 586-620.*

[Su] V.N.Sudakov, *Proceedings of the Steklov Institute (1979) vol. 141.*

[Ta] L.Tartar, *The compensated compactness method applied to systems of conservation laws. Systems of nonlinear PDE, NATO ASI series, Reidel,Dordecht, 1983.*

[Ve] A.M. Vershik, *Russian Math. Surveys 25 (1970) vol. 5, 117-124.*

[Yo] L.C.Young, *Lectures on the calculus of variations. Chelsea,New York, 1980.*

[ZM] Y.D. Zheng, A. Majda, *Existence of global weak solutions to one-component Vlasov-Poisson and Fokker-Planck-Poisson systems in one space dimension with measures as initial data, Comm. Pure Appl. Math. 47 (1994) 1365-1401.*