CONNECTIONS BETWEEN OPTIMAL TRANSPORT, COMBINATORIAL OPTIMIZATION AND HYDRODYNAMICS

Yann Brenier[1]

Abstract. We discuss a new connection between combinatorial optimization and optimal transport theory through the analysis of a variational problem coming from mathematical Fluid Mechanics. At a discrete level, this minimization problem corresponds to a quadratic assignment problem, which belongs to the NP class of combinatorial optimization. Our analysis is focused on the study of a suitable gradient flow for which we establish the global existence of dissipative solutions which are unique when smooth.

1. Introduction

There are well-established connections between combinatorial optimization, optimal transport theory and Hydrodynamics. For instance, the famous earth moving problem of Monge ("problème des déblais et des remblais" [19]) corresponds, after discretization, to the "linear assignment problem", which is an easy "P" problem in combinatorial optimization. The Monge problem is clearly related to Continuum Mechanics and its convex relaxation, known as the Monge−Kantorovich problem [19], admits a formulation in terms of Fluid Mechanics [2, 6], which corresponds to a quite trivial physical model: a potential inviscid flow without pressure! So, it is natural to look for more sophisticated interactions between these three fields of combinatorial optimization, optimal transport theory and Fluid Mechanics. In the present paper, we consider the following variational problem: let D be a smooth domain in $\mathbb{R}^d$ and a φ_0 a smooth function on D valued in $\mathbb{R}^N$. We denote by λ the "law" of φ_0 over $\mathbb{R}^N$, defined by

$$\int_D F(\varphi_0(x))\mathrm{d}x = \int_{\mathbb{R}^N} F(y)\lambda(\mathrm{d}y), \ \ \forall F \in C_c^0\left(\mathbb{R}^N\right). \tag{1.1}$$

We want to minimize the Dirichlet integral

$$\mathcal{E}[\varphi] = \frac{1}{2}\int_D |\nabla\varphi(x)|^2\mathrm{d}x \tag{1.2}$$

among all functions φ from D to $\mathbb{R}^N$, with law λ. This problem is closely related to the (stationary) solutions of the famous (and highly non-trivial) Euler equations of incompressible Fluid Mechanics, in the case $d = 2, N = 1$.

[1] CNRS (CMLS, UMR7640), École Polytechnique, 91128, Palaiseau, France. `yann.brenier@polytechnique.edu`

At the discrete level, it corresponds to a "quadratic assignment problem", a famous "NP problem" in combinatorial optimization.

In the present paper, we do not address this difficult non-convex minimization problem, but rather introduce and analyze a "gradient flow" equation for it, namely

$$\partial_t \varphi + \nabla \cdot (\varphi v) = 0, \quad (-\triangle)^m v = -\mathbb{P}\nabla \cdot (\nabla\varphi \otimes \nabla\varphi), \tag{1.3}$$

where $\mathbb{P}$ denotes the L^2 projection onto divergence-free vector fields and $m = 0$ or $m = 1$, with suitable boundary conditions.

Then, combining some ideas of P.-L. Lions (for the Euler equations) and Ambrosio−Gigli-Savaré (for the heat equation), we provide for the initial value problem a concept of generalized "dissipative" solutions which always exist globally in time and are unique whenever they are smooth.

Remark 1.1. After the submission of this paper, we got informed by Filippo Santambrogio of reference [3] (and also [12,14]). In that work about "area preserving mappings of minimal distortion", Angenent *et al.* motivated by medical imaging problems, had already proposed a gradient flow approach to our minimization problem, at least in the case $d = N = 2$. In their approach, the gradient flow is computed by differential geometry methods and its analysis has been postponed to a future publication. Our derivation is equivalent (but done in "Eulerian coordinates" instead). As a matter of fact, we have limited our analysis, in the rest of the present paper, to the case $N = 1$, but our methodology extends without difficulty to the multidimensional case $N > 1$. Let us finally mention the somewhat related work of Gay−Balmaz and Holm [11] in their approach of partly dissipative models in Fluid Mechanics.

2. WELL-KNOWN CONNECTIONS BETWEEN OPTIMAL TRANSPORT THEORY, HYDRODYNAMICS AND COMBINATORIAL OPTIMIZATION

2.1. The Monge−Kantorovich distance in optimal transport theory

The (quadratic) Monge−Kantorovich (MK^2) distance (very often called "Wasserstein" distance in optimal transport theory and also called Tanaka distance in kinetic theory) can be defined in terms of probability measures and random variables as:

$$d_{MK^2}(\mu, \nu) = \inf \left\{ \sqrt{E(|X - Y|^2)}, \quad \text{law}(X) = \mu, \ \text{law}(Y) = \nu \right\} \tag{2.1}$$

where μ and ν are probability measures (with finite second moments) defined on the Euclidean space $\mathbb{R}^d$, X and Y denotes random variables valued in $\mathbb{R}^d$, $|\cdot|$ is the Euclidean norm and E denotes the expected value.

2.2. Hydrodynamic interpretation of the $\mathsf{MK^2}$ distance

Using the so-called "Benamou−Brenier formula" or the "Otto calculus" [6,17], we may express the MK^2 distance in hydrodynamic terms. More precisely, at least in the case when μ and ν are absolutely continuous with respect to the Lebesgue measure, we may write

$$d_{MK^2}^2(\mu, \nu) = \inf \ \int_0^1 \int_{\mathbb{R}^d} |v(t,x)|^2 \rho(t,x) \mathrm{d}t \mathrm{d}x \tag{2.2}$$

where the infimum is taken over all density and velocity fields

$$(t, x) \in [0, 1] \times \mathbb{R}^d \to (\rho(t, x), v(t, x)) \in \mathbb{R}_+ \times \mathbb{R}^d$$

subject to the continuity equation

$$\partial_t \rho + \nabla_x \cdot (\rho v) = 0 \tag{2.3}$$

and the time-boundary conditions

$$\rho(t=0,x)\mathrm{d}x = \mu(\mathrm{d}x), \quad \rho(t=1,x)\mathrm{d}x = \nu(\mathrm{d}x). \tag{2.4}$$

The (formal) optimality equations read

$$v(t,x) = \nabla_x \theta(t,x), \quad \partial_t \theta + \frac{1}{2}|\nabla_x \theta|^2 = 0, \tag{2.5}$$

and describe a potential, inviscid, pressure-less gas, sometimes called "dust" (in cosmology in particular), which is one of the most trivial models of fluids.

2.3. $\mathsf{MK^2}$ distance and combinatorial optimization

Given two discrete probability measures on $\mathbb{R}^d$

$$\mu = \sum_{i=1,N} \delta_{A_i}, \quad \nu = \sum_{j=1,N} \delta_{B_j}, \tag{2.6}$$

we easily check that

$$d^2_{MK^2}(\mu,\nu) = \inf_{\text{law}(X)=\mu,\text{law}(Y)=\nu} E(|X-Y|^2) = \inf_{\sigma \in \mathcal{S}_N} \sum_{i=1,N} \frac{|A_i - B_{\sigma_i}|^2}{N}, \tag{2.7}$$

where $\sigma \in \mathcal{S}_N$ denotes the set of all permutations of the first N integers. Thus, computing the MK^2 distance between two discrete measures is equivalent to solving the so-called "linear assignment problem" (LAP):

$$\inf_{\sigma \in \mathcal{S}_N} \sum_{i=1}^{N} c(i,\sigma_i), \tag{2.8}$$

in the special case when the "cost matrix" c has geometric contents

$$c(i,j) = |A_i - B_j|^2. \tag{2.9}$$

In full generality, the LAP is one of the simplest combinatorial optimization problems, with complexity $O(N^3)$ [5].

3. NP COMBINATORIAL OPTIMIZATION PROBLEMS AND HYDRODYNAMICS

There are much more challenging problems in combinatorial optimization, such as the (NP) "quadratic assignment problem" (which includes the famous traveling salesman problem).

Given two $N \times N$ matrices γ and c, with coefficients ≥ 0, solve:

$$(QAP) \quad \inf_{\sigma \in \mathcal{S}_N} \sum_{i,j=1,N} c(\sigma_i,\sigma_j)\gamma(i,j). \tag{3.1}$$

The QAP is useful in computer vision [16]. Some continuous versions of the QAP are related to recent works in geometric and functional analysis [18]. It turns out that the QAP can also be related to Hydrodynamics, as we are going to see.

3.1. A minimization problem in hydrodynamics

The minimization problem of the Dirichlet integral (1.2) under the law constraint (1.1), discussed in the introduction, seems to go back to Lord Kelvin and has been frequently studied since (see, for instance [7,8]). It reads

$$\inf\left\{ \frac{1}{2}\int_D |\nabla\varphi(x)|^2 dx, \quad \mathrm{Law}(\varphi) = \mathrm{Law}(\varphi_0) = \lambda \right\} \tag{3.2}$$

and can be immediately rephrased as a saddle-point problem:

$$\inf_\varphi \sup_{F:\mathbb{R}\to\mathbb{R}} \quad \frac{1}{2}\int_D |\nabla\varphi(x)|^2 dx + \int_D F(\varphi(x)) dx - \int_\mathbb{R} F(y)\lambda(dy). \tag{3.3}$$

Optimal solutions are formally solutions to

$$- \triangle\,\varphi + F'(\varphi) = 0, \tag{3.4}$$

for some function $F : \mathbb{R} \to \mathbb{R}$, and, in 2d, are just stationary solutions to the Euler equations of incompressible fluids [4,15]. (More precisely, φ is the stream-function of a stationary two-dimensional incompressible inviscid fluid).

3.2. The discrete version of the hydrodynamic problem is a QAP

Let us discretize the domain D with a lattice of N vertices $A_1, \ldots, A_N$ and define coefficients $\gamma(i,j) \geq 0$ so that the Dirichlet integral of a function φ can be approximated as follows:

$$\int_D |\nabla\varphi(x)|^2 dx \ \sim \ \sum_{i,j=1}^N \gamma(i,j)|\varphi(A_i) - \varphi(A_j)|^2.$$

At the discrete level, we may say that φ and φ_0 have the same (discrete law) whenever

$$\varphi(A_i) = \varphi_0(A_{\sigma_i}), \quad i = 1, \ldots, N,$$

for some permutation $\sigma \in \mathcal{S}_N$. Thus, the discrete version of (3.2) reads:

Find a permutation σ that achieves

$$\inf_\sigma \ \sum_{i,j=1}^N c(\sigma_i, \sigma_j)\gamma(i,j) \tag{3.5}$$

with

$$c(i,j) = |\varphi_0(A_i) - \varphi_0(A_j)|^2. \tag{3.6}$$

So we have clearly obtained a particular case of QAP (3.1).

4. A "GRADIENT-FLOW" APPROACH TO THE HYDRODYNAMIC PROBLEM

To address problem (3.2), it is natural to use a "gradient flow" approach involving a time dependent function $\varphi_t(x)$ starting from $\varphi_0(x)$ at $t = 0$. Hopefully, as $t \to +\infty$, φ_t will reach a solution to our minimization problem. (Nevertheless, there is no guarantee at all that we can reach a global minimizer, not even a local minimizer. This is clearly a widely open problem, due, in particular, to the non-convexity of the minimization problem). A canonical way of preserving the law λ of $\varphi(t, \cdot)$ during the evolution is the transport of φ by a (sufficiently) smooth time-dependent divergence-free velocity field $v = v_t(x) \in \mathbb{R}^d$, parallel to ∂D, according to

$$\partial_t\varphi_t + \nabla\cdot(v_t\varphi_t) = 0, \ \nabla\cdot v_t = 0, \ v_t//\partial D. \tag{4.1}$$

Indeed, we easily get:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_D F(\varphi_t(x))\mathrm{d}x = -\int_D F'(\varphi_t(x))\nabla \cdot (v_t(x)\varphi_t(x))\mathrm{d}x$$

$$= -\int_D v_t(x) \cdot \nabla(F(\varphi_t(x)))\mathrm{d}x = 0$$

(since v is divergence-free), for all smooth bounded functions F. Loosely speaking, the vector field v should be interpreted as a kind of "tangent vector" along the "orbit" of all φ sharing the same law λ as φ_0.

From the analysis viewpoint, according to the DiPerna−Lions theory on ODEs [10], for the law λ to be preserved, there is no need for v to be very smooth and it is just enough that the space derivatives of v are Lebesgue integrable functions (or even bounded Borel measures, according to Ambrosio [1]):

$$\int_0^T \int_D |\nabla v_t(x)|\mathrm{d}x\mathrm{d}t < +\infty, \quad \forall T > 0.$$

In that situation, the solution of (4.1) is just (implicitly) given by $\varphi_t(\xi_t(x)) = \varphi_0(x)$, where ξ is the unique time-dependent family of almost-everywhere one-to-one volume-preserving Borel maps of D generated by v through:

$$\partial_t \xi_t(x) = v_t(\xi_t(x)), \ \xi_0(x) = x.$$

This is enough to preserve the law of φ_t as t evolves. (Notice, however, that, if v is smoother, say of class C^1, the maps ξ_t become orientation preserving diffeomorphisms, which makes the evolution much more constrained and, as a consequence, further reduces the chance to reach a minimizer as $t \to +\infty$, unless the initial condition φ_0 is carefully chosen).

In order to get a "gradient flow", we also need a quadratic form (or, more generally, a convex functional, which would then rather correspond to a "Finslerian flow") acting on the "tangent vector" v. For this purpose, let us first denote by Sol(D) the Hilbert space of all square-integrable divergence free vector fields on D and parallel to ∂D, with L^2 norm and inner-product respectively denoted by $|| \cdot ||$ and $((\cdot, \cdot))$. Next, let us fix a lower semi-continuous convex functional $K : a \in \text{Sol(D)} \to K[a] \in [0, +\infty]$. We assume that at each smooth vector-field ω in Sol(D),

(i) K is finite;
(ii) its subgradient has a unique element $K'[\omega] \in L^2(D, \mathbb{R}^d)$;
(iii) there is $\epsilon_K[\omega] > 0$ such that, the "relative entropy" of K controls the L^2 distance:

$$\eta_K[v, \omega] = K[v] - K[\omega] - ((K'[\omega], v - \omega)) \geq \epsilon_k[\omega]||v - \omega||^2, \quad \forall v \in \text{Sol(D)}. \tag{4.2}$$

The simplest example is of course

$$K[v] = \frac{1}{2} \int_D |v(x)|^2 \mathrm{d}x. \tag{4.3}$$

Then, we are given a "functional" $\varphi \in E \to \mathcal{E}[\varphi] \in \mathbb{R}$ on a suitable function space E, the canonical example for us being the Dirichlet integral (1.2) over the Sobolev space $E = H_0^1(D)$. When we evolve φ according to (4.1), we formally get

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[\varphi_t] = \int_D \mathcal{E}'[\varphi_t](x)\partial_t\varphi_t(x)\mathrm{d}x = -\int_D \mathcal{E}'[\varphi_t](x)\nabla \cdot (v_t(x)\varphi_t(x))\mathrm{d}x,$$

where we denote by $\mathcal{E}'$ the gradient of $\mathcal{E}$ with respect to the L^2 metric. (In the case of the Dirichlet integral, $\mathcal{E}'[\varphi] = -\triangle \varphi$.) Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[\varphi_t] = -\int_D \mathcal{E}'[\varphi_t](x)\nabla\varphi_t(x) \cdot v_t(x)\mathrm{d}x \tag{4.4}$$

(using that v_t is divergence-free).

Then, we may write (4.4) as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[\varphi_t] = -((G_t, v_t)), \tag{4.5}$$

$$G_t = \mathbb{P}(\mathcal{E}'[\varphi_t]\nabla\varphi_t), \tag{4.6}$$

where we denote by $\mathbb{P}$ the L^2 projection onto Sol(D). Thus, denoting by K^* the Legendre−Fenchel transform $K^*[g] = \sup_w((g, w)) - K[w]$, we get:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[\varphi_t] + K[v_t] + K^*[G_t] = K[v_t] + K^*[G_t] - ((G_t, v_t))$$

where by definition of the Legendre−Fenchel transform, the right-hand side is always nonnegative and vanishes if and only if

$$v_t = K^{*\prime}[G_t] \tag{4.7}$$

(for instance, in case (4.3), $v_t = G_t$). Equation (4.7) precisely is the "closure equation" we need to define the "gradient flow" of $\mathcal{E}$ with respect to the evolution equation (4.1) with "metric" K. (This way, we closely follow [9], in the spirit of [2].) As just seen, this closure equation is equivalent to the differential inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[\varphi_t] + K[v_t] + K^*[G_t] \le 0,$$

or, using the definition of K^* as the Legendre−Fenchel transform of K,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}[\varphi_t] + K[v_t] + ((G_t, z_t)) - K[z_t] \le 0, \ \forall \ z_t \in \text{Sol(D)}. \tag{4.8}$$

So, our gradient flow is now defined by combining transport equation (4.1), definition (4.6), and either the closure equation (4.7) or the variational inequality (4.8), which are formally equivalent.

5. **The gradient flow equation**

From now on, let us consider, for simplicity, the case of the periodic cube $D = \mathbb{T}^d$, instead of a bounded domain of $\mathbb{R}^d$. Accordingly, all functions (φ_t, v_t, $etc.\dots$) to be considered will be of zero mean in $x \in D$. We also concentrate on the case when:

(i) $\mathcal{E}$ is the Dirichlet integral (1.2);
(ii) K is the Sobolev (semi-)norm of order m, $m \in \{0, 1, 2, \dots\}$,

$$K[v] = \frac{1}{2}\int_D |\nabla^m v(x)|^2 \mathrm{d}x, \tag{5.1}$$

which should be understood, when $m > 0$, as the H^m Sobolev semi-norm of v when it makes sense and $+\infty$ otherwise.

Then, $\mathcal{E}' = -\triangle$, $K^{*\prime} = (-\triangle)^{-m}$. Also notice that the "relative entropy" reads, for each pair (v, ω) in Sol(D) with ω smooth,

$$\eta_K[v, \omega] = K[v] - K[\omega] - ((K'[\omega], v - \omega)) = K[v - \omega] \ge c||v - \omega||^2,$$

where $c > 0$ depends only on m and d, by Poincaré's inequality on the periodic cube. So, equations (4.7) and (4.6) respectively become

$$v_t = (-\triangle)^{-m} G_t ,$$

$$G_t = \mathbb{P}(\mathcal{E}'[\varphi_t] \nabla \varphi_t) = \mathbb{P}(-\triangle \varphi_t \nabla \varphi_t) = \mathbb{P}\left[-\nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t) + \frac{1}{2} \nabla(|\nabla \varphi_t|^2) \right]$$
$$= -\mathbb{P} \nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t)$$

(since $\mathbb{P}$ is the L^2 projection onto Sol(D) and, therefore, cancels any gradient). Thus

$$v_t = -(-\triangle)^{-m} \mathbb{P} \nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t). \tag{5.2}$$

Similarly, (4.8) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\nabla \varphi_t\|^2 + 2K[v_t] + ((\nabla \varphi_t \otimes \nabla \varphi_t, \nabla z_t + \nabla z_t^T)) \leq 2K[z_t], \tag{5.3}$$

for all smooth $z_t \in$ Sol(D).

Finally, the gradient-flow equation reads:

$$\partial_t \varphi_t + \nabla \cdot (\varphi_t v_t) = 0, \quad (-\triangle)^m v_t = -\mathbb{P} \nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t). \tag{5.4}$$

5.1. Physical interpretation of the GF equation

Physically speaking, the GF (gradient flow) equation (5.4) in the case $m = 1$ corresponds to the "Stokes flow"

$$\partial_t \varphi_t + \nabla \cdot (\varphi_t v_t) = 0, \quad -\triangle v_t = -\mathbb{P} \nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t) \tag{5.5}$$

of an electrically charged incompressible fluid (v and φ being the velocity and the electric potential), while the case $m = 0$ rather corresponds to a "Darcy flow"

$$\partial_t \varphi_t + \nabla \cdot (\varphi_t v_t) = 0, \quad v_t = -\mathbb{P} \nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t). \tag{5.6}$$

Observe, in both cases, that the electric potential φ is not external but, rather, coupled to the velocity field. As a matter of fact, both models can be seen as two different dissipative versions (one for very viscous fluids and the other one for flows in porous media) of the "ideal Electrohydrodynamics" equations

$$\partial_t \varphi_t + \nabla \cdot (\varphi_t v_t) = 0, \quad \partial_t v_t + \nabla \cdot (v_t \otimes v_t) = -\mathbb{P} \nabla \cdot (\nabla \varphi_t \otimes \nabla \varphi_t), \tag{5.7}$$

which are the electric counterpart of the better known equations of 3D ideal Magnetohydrodynamics:

$$\partial_t B_t + \nabla \times (B_t \times v_t) = 0, \quad \partial_t v_t + \nabla \cdot (v_t \otimes v_t) = \mathbb{P} \nabla \cdot (B_t \otimes B_t). \tag{5.8}$$

6. Analysis of the gradient flow equation

The last part of this article is devoted to the analysis of the gradient flow equation (5.4). For that purpose, we closely follow the ideas and concepts of our recent work [9].

6.1. A concept of "dissipative solutions"

From the analysis viewpoint, we ignore whether or not gradient-flow (GF) equation (5.4), namely

$$\partial_t \varphi_t + \nabla \cdot (\varphi_t v_t) = 0, \quad (-\triangle)^m v_t = -\mathbb{P}\nabla \cdot (\nabla\varphi_t \otimes \nabla\varphi_t),$$

is locally well-posed in any space of smooth functions (unless $m > d/2+1$). The global existence of weak solutions can be expected for the Stokes version (5.5) (with φ a priori in $L_t^\infty(H_x^1)$ and v "almost" in $L_t^\infty(W_x^{1,1})$), while such a result looks out of reach in the case of the "Darcy" version (5.6).

Anyway, we prefer a much more "robust" concept of solutions, that we call "dissipative" after [9], somewhat in the spirit of Lions' dissipative solutions to the Euler's equations [13] and following some ideas of the analysis of the linear heat equations for general measured metric spaces by Ambrosio *et al.* [2]. We keep transport equation (4.1) and integrate (5.3) on $[0,t]$, for all $t \geq 0$, with a suitable exponential weight, which leads to:

$$\int_0^t \{2K[v_s] + ((\nabla\varphi_s \otimes \nabla\varphi_s, rI_d + \nabla z_s + \nabla z_s^T)) - 2K[z_s]\}\mathrm{e}^{-sr}\mathrm{d}s \tag{6.1}$$

$$+||\nabla\varphi_t||^2\mathrm{e}^{-tr} \leq ||\nabla\varphi_0||^2, \quad \text{for every smooth field } t \to z_t \in \mathrm{Sol(D)}.$$

Here $r \geq 0$ is a constant, depending on z, chosen so that

$$\forall (t,x), \quad rI_d + \nabla z_t(x) + \nabla z_t(x)^T \geq 0, \quad \text{in the sense of symmetric matrices}, \tag{6.2}$$

in order to be sure that inequality (6.1) only involves convex functionals of (φ, v). From the functional analysis viewpoint, it is natural to consider solutions $(B_t = \nabla\varphi_t, v_t, t \in [0,T])$, for each fixed $T > 0$, in the space

$$C_w^0([0,T], L^2(D,\mathbb{R}^d)) \times L^2([0,T], \mathrm{Sol}(D)),$$

where $C_w^0(L^2)$ just means continuity in time with respect to the weak topology of L^2.

6.2. Uniqueness of smooth solutions among dissipative solutions

Theorem 6.1. *Assume $D = (\mathbb{R}/\mathbb{Z})^d$ and define K by (5.1), namely*

$$K[v] = \frac{1}{2}\int_D |\nabla^m v(x)|^2 \mathrm{d}x,$$

with "relative entropy"

$$\eta_K[a,b] = K[a] - K[b] - ((K'[b], a-b)) = \frac{1}{2}\int_D |\nabla^m(a-b)(x)|^2\mathrm{d}x. \tag{6.3}$$

Let us fix $T > 0$ and consider

$$(B_t = \nabla\varphi_t, v_t, t \in [0,T]) \in C_w^0([0,T], L^2(D,\mathbb{R}^2)) \times L^2([0,T], \mathrm{Sol}(D)),$$

a dissipative solution of the GF equation (5.4) up to time T, in the sense of (4.1, 6.1, 6.2). Let $(\beta_t = \nabla\psi_t, \omega_t, t \in [0,T])$ be any pair of smooth functions with ω valued in $\mathrm{Sol(D)}$. Then there is a constant C depending only on K and the spatial Lipschitz constant of (β, ω), up to time T, so that, for all $t \in [0,T]$,

$$||B_t - \beta_t||^2 + \int_0^t \mathrm{e}^{(t-s)C}\{\eta_K[v_s, \omega_s]\mathrm{d}s - 2J_s^L\}\mathrm{d}s \leq ||B_0 - \beta_0||^2\mathrm{e}^{tC}$$

$$J_t^L = -((B_t - \beta_t, \nabla(\omega_t \cdot \beta_t) + \partial_t\beta_t)) - ((\mathbb{P}\nabla \cdot (\beta_t \otimes \beta_t) + K'[\omega_t], v_t - \omega_t)). \tag{6.4}$$

In particular, J_t^L exactly vanishes as $(\beta = \nabla\psi, \omega)$ is a smooth solution to the GF equation (5.4), *namely*

$$\partial_t \beta_t + \nabla(\omega_t \cdot \beta_t) = 0, \quad K'[\omega_t] = -\mathbb{P}\nabla(\beta_t \otimes \beta_t),$$

in which case

$$||B_t - \beta_t||^2 + \int_0^t e^{(t-s)C} \eta_K[v_s, \omega_s] ds \ \le ||B_0 - \beta_0||^2 e^{tC}. \tag{6.5}$$

This implies the uniqueness of smooth solutions among all dissipative solutions, for any given prescribed smooth initial condition.

6.3. Global existence of dissipative solutions

In the spirit of [9], at least in the case: $D = \mathbb{T}^d$,

$$\mathcal{E}[\varphi] = \frac{1}{2} \int_{\mathbb{T}^d} |\nabla\varphi(x)|^2 dx, \quad K[v] = \frac{1}{2} \int_{\mathbb{T}^d} |(\nabla)^m v(x)|^2 dx \quad (m = 0, 1, 2, \ldots),$$

it is fairly easy to establish, for the "dissipative" formulation (4.1), (6.1), (6.2) of (4.1), (4.7) and for each initial condition φ_0 with finite Dirichlet integral, the existence of a global solution $(B = \nabla\varphi, v)$ in $C_w^0(\mathbb{R}_+, L^2(D, \mathbb{R}^d)) \times L^2(\mathbb{R}_+, \mathrm{Sol}(D))$. Without entering into details, let us sketch the proof. We approximate B_0 strongly in L^2 by some smooth field $B_0^\epsilon = \nabla\varphi_0^\epsilon$ and mollify K by substituting for it

$$K^{M,\epsilon}(v) = K(v) + \epsilon ||\nabla^M v||^2$$

with M sufficiently large $(M > d/2 + 1)$ and $\epsilon > 0$. In this case, we get, M and ϵ being fixed, a uniform *a priori* bound for v in $L^2([0, T], C^1(D))$, which is enough to solve transport equation (4.1) in the classical framework of the Cauchy−Lipschitz theory of ODEs. Then, we get a smooth approximate solution $(B^\epsilon = \nabla\varphi^\epsilon, v^\epsilon)$ satisfying transport equation (4.1), *i.e.*,

$$\partial_t \varphi_t^\epsilon + \nabla \cdot (\varphi_t^\epsilon v_t^\epsilon) = 0,$$

together with (6.1,6.2), namely

$$\int_0^t \{2K^{M,\epsilon}[v_s^\epsilon] + ((B_s^\epsilon \otimes B_s^\epsilon, rI_d + \nabla z_s + \nabla z_s^T)) - 2K^{M,\epsilon}[z_s]\} e^{-sr} ds$$

$$+||B_t^\epsilon||^2 e^{-tr} \le ||B_0^\epsilon||^2, \quad \text{for every smooth field } t \to z_t \in \mathrm{Sol(D)}$$

satisfying (6.2), and, in particular (for $z = 0$)

$$\int_0^t 2K^{M,\epsilon}[v_s^\epsilon] ds + ||B_t^\epsilon||^2 \le ||B_0^\epsilon||^2.$$

We get enough compactness for the approximate solutions to get a limit (B, v) in space

$$C_w^0(\mathbb{R}_+, L^2(D, \mathbb{R})) \times L^2(\mathbb{R}_+, \mathrm{Sol}(D)),$$

and pass to the limit in the transport equation (since $B^\epsilon = \nabla\varphi^\epsilon$). Finally, by lower semi-continuity, we may pass to the limit in the dissipation inequality and obtain (6.1) and (6.2), which concludes the (sketch of) proof.

Observe that, for $m \ge 1$, the L^2 norm of ∇v_t is square integrable in time. This implies by DiPerna−Lions ODE theory (see [13]), as already discussed, that the law of φ_t stays unchanged during the evolution by (4.1) (but, unless $m > 1 + d/2$, not necessarily its topology, which is of some interest in view of the minimization problem (3.2) we started with). However, unless $m > 1 + d/2$, it is unclear to us that (4.1) and (4.7) even admits local smooth solutions.

APPENDIX A. PROOF OF THEOREM 6.1

Choose $r \geq 0$ such that ω satisfies (6.2), namely

$$\forall (t,x), \quad rI_d + \nabla \omega_t(x) + \nabla \omega_t(x)^T \geq 0, \quad \text{in the sense of symmetric matrices.}$$

Since $(B = \nabla \varphi, v)$ is a dissipative solution, we get, by setting $z = \omega$ in definition (6.1),

$$\int_0^t \{2K[v_s] + ((B_s \otimes B_s, rI_d + \nabla \omega_s + \nabla \omega_s^T)) - 2K[\omega_s]\} e^{-sr} ds \tag{A.1}$$

$$+ ||B_t||^2 e^{-tr} \leq ||B_0||^2.$$

Let us now introduce for each $t \in [0, T]$

$$N_t = ||B_0||^2 e^{rt} - \int_0^t \{2K[v_s] + ((B_s \otimes B_s, rI_d + \nabla \omega_s + \nabla \omega_s^T)) - 2K[\omega_s]\} e^{r(t-s)} ds \tag{A.2}$$

so that

$$N_t \geq ||B_t||^2, \forall t \in [0, T].$$

By definition (A.2) of N_t, we have

$$\left(\frac{d}{dt} - r\right) N_t = -2K[v_t] - \left((B_t \otimes B_t, rI_d + \nabla \omega_t + \nabla \omega_t^T)\right) + 2K[\omega_t]$$

and, therefore,

$$\frac{d}{dt} N_t = r(N_t - ||B_t||^2) - 2K[v_t] - \left((B_t \otimes B_t, \nabla \omega_t + \nabla \omega_t^T)\right) + 2K[\omega_t] \tag{A.3}$$

(in the distributional sense and also for a.e. $t \in [0, T]$).
We now want to estimate

$$e_t = ||B_t - \beta_t||^2 + (N_t - ||B_t||^2) = N_t - 2((B_t, \beta_t)) + ||\beta_t||^2, \ \forall t \in [0, T], \tag{A.4}$$

where $\beta_t = \nabla \psi_t$. Since $(B = \nabla \varphi, v)$ is a dissipative solution, it solves transport equation (4.1) which implies

$$\partial_t B_t + \nabla(B_t \cdot v_t) = 0,$$

after derivation in x. Thus

$$\frac{d}{dt}((B_t, \beta_t)) = \int B_{ti}(\beta_{ti,t} + v_{ti}\beta_{tj,j})$$

(where we use notations $\beta_{ti,j} = \partial_j(\beta_t)_i$, $etc.\ldots$ and skip summations on repeated indices $i, j\ldots$).
   Using (A.3) and definition (A.4), we deduce

$$\frac{d}{dt} e_t = r(N_t - ||B_t||^2) - 2K[v_t] - ((B_t \otimes B_t, \nabla \omega_t + \nabla \omega_t^T)) + 2K[\omega_t]$$

$$+ \int 2(\beta - B)_{ti}\beta_{ti,t} - 2B_{ti}v_{ti}\beta_{tj,j}.$$

Thus

$$\frac{d}{dt} e_t = r(N_t - ||B_t||^2) - 2K[v_t] + 2K[\omega_t] + J_t$$

where

$$J_t = \int -B_{ti}B_{tj}(\omega_{ti,j} + \omega_{tj,i}) + 2(\beta - B)_{ti}\beta_{ti,t} - 2B_{ti}v_{ti}\beta_{tj,j}.$$

Denoting the "relative entropy" of K by $\eta_K[a,b] = K[a] - K[b] - ((K'[b], a - b))$, we have obtained

$$\frac{\mathrm{d}}{\mathrm{d}t}e_t + 2\eta_K[v_t, \omega_t] = r(N_t - ||B_t||^2) + J_t - 2((K'[\omega_t], v_t - \omega_t)). \tag{A.5}$$

We may write

$$J_t = J_t^Q + J_t^{L1} + J_t^{L2} + J_t^C$$

where J_t^Q, J_t^{L1}, J_t^{L2}, J_t^C are respectively quadratic, linear, linear, and constant with respect to $B - \beta$ and $v - \omega$, with coefficient depending only on ω, β:

$$J_t^Q = \int -(B-\beta)_{ti}(B-\beta)_{tj}(\omega_{ti,j} + \omega_{tj,i}) - 2(B-\beta)_{ti}(v-\omega)_{ti}\beta_{tj,j}$$

$$J_t^{L1} = \int 2(B-\beta)_{ti}[-\beta_{tj}(\omega_{ti,j} + \omega_{tj,i}) - \beta_{ti,t} - \omega_{ti}\beta_{tj,j}]$$

$$J_t^{L2} = -\int 2(v-\omega)_{ti}\beta_{ti}\beta_{tj,j}$$

$$J_t^C = \int [-\beta_{ti}\beta_{tj}(\omega_{ti,j} + \omega_{tj,i}) - 2\beta_{ti}\omega_{ti}\beta_{tj,j}].$$

Let us reorganize these four terms. By integration by part of its first term, we see that $J_t^C = 0$, using that β is a gradient and ω is divergence-free. More precisely

$$-\int \beta_{ti}\beta_{tj}(\omega_{ti,j} + \omega_{tj,i}) = -2\int \beta_{ti}\beta_{tj}\omega_{ti,j} = 2\int \beta_{ti,j}\beta_{tj}\omega_{ti} + 2\int \beta_{ti}\beta_{tj,j}\omega_{ti}$$

$$= \int (\beta_{tj}\beta_{tj})_{,i}\omega_{ti} + 2\int \beta_{ti}\beta_{tj,j}\omega_{ti} = 0 + 2\int \beta_{ti}\beta_{tj,j}\omega_{ti}.$$

Using that $v_t - \omega_t$ is divergence-free while β is a gradient, we immediately get

$$J_t^{L2} = -2((\mathbb{P}\nabla \cdot (\beta_t \otimes \beta_t), v_t - \omega_t)).$$

Next, we find

$$J_t^{L1} = -2((B_t - \beta_t, \nabla(\omega_t \cdot \beta) + \partial_t \beta_t)).$$

Indeed, since

$$J_t^{L1} = \int 2(B-\beta)_{ti}[-\beta_{tj}(\omega_{ti,j} + \omega_{tj,i}) - \beta_{ti,t} - \omega_{ti}\beta_{tj,j}]$$

we get

$$J_t^{L1} + 2((B_t - \beta_t, \nabla(\omega_t \cdot \beta) + \partial_t \beta_t)) = \int 2(B-\beta)_{ti}[-\beta_{tj}(\omega_{ti,j} + \omega_{tj,i}) + (\omega_{tj}\beta_{tj})_{,i} - \omega_{ti}\beta_{tj,j}]$$

$$= \int 2(B-\beta)_{ti}[-(\beta_{tj}\omega_{ti})_{,j} - \beta_{tj}\omega_{tj,i} + \omega_{tj,i}\beta_{tj} + \omega_{tj}\beta_{tj,i}] = \int 2(B-\beta)_{ti}[-(\beta_{tj}\omega_{ti})_{,j} + \omega_{tj}\beta_{tj,i}]$$

$$= \int 2(B-\beta)_{ti}[-(\beta_{tj}\omega_{ti})_{,j} + (\beta_{ti}\omega_{tj})_{,j} - \beta_{ti}\omega_{tj,j}] = 0$$

(since ω is divergence-free while $B_t - \beta_t$ and β_t are gradients). Next, since

$$J_t^Q = \int -(B-\beta)_{ti}(B-\beta)_{tj}(\omega_{ti,j} + \omega_{tj,i}) - 2(B-\beta)_{ti}(v-\omega)_{ti}\beta_{tj,j}$$

we may find, for any fixed $\epsilon > 0$, a constant C_ϵ (depending on the spatial Lipschitz constant of (β, ω)) such that

$$J_t^Q \leq \epsilon||v_t - \omega_t||^2 + C_\epsilon||B_t - \beta_t||^2.$$

Using (6.3), we may choose ϵ small enough so that

$$\epsilon||v_t - \omega_t||^2 \leq \eta_K[v_t, \omega_t].$$

So, we get from (A.5)

$$\frac{\mathrm{d}}{\mathrm{d}t}e_t + \eta_K[v_t, \omega_t] \leq r(N_t - ||B_t||^2) + C_\epsilon||B_t - \beta_t||^2 + 2J_t^L \tag{A.6}$$

where

$$J_t^L = -((B_t - \beta_t, \nabla(\omega_t \cdot \beta) + \partial_t\beta_t)) - ((\mathbb{P}\nabla \cdot (\beta_t \otimes \beta_t) + K'[\omega_t], v_t - \omega_t)).$$

By definition (A.4) of e_t, namely $e_t = ||B_t - \beta_t||^2 + (N_t - ||B_t||^2)$, we have obtained

$$\frac{\mathrm{d}}{\mathrm{d}t}e_t + \eta_K[v_t, \omega_t] \leq Ce_t + 2J_t^L$$

for a constant C depending only on β, ω and K. By integration we deduce

$$e_t + \int_0^t \mathrm{e}^{(t-s)C}\eta_K[v_s, \omega_s]\mathrm{d}s \leq e_0\mathrm{e}^{tC} + 2\int_0^t \mathrm{e}^{(t-s)C}J_s^L\mathrm{d}s.$$

Next, let us remind that $e_t \geq ||B_t - \beta_t||^2$ with equality for $t = 0$ (since $N_t \geq ||B_t||^2$ with equality at $t = 0$). Thus, we have finally obtained

$$||B_t - \beta_t||^2 + \int_0^t \mathrm{e}^{(t-s)C}\eta_K[v_s, \omega_s]\mathrm{d}s \leq ||B_0 - \beta_0||^2\mathrm{e}^{tC} + 2\int_0^t \mathrm{e}^{(t-s)C}J_s^L\mathrm{d}s$$

with

$$J_t^L = -((B_t - \beta_t, \nabla(\omega_t \cdot \beta) + \partial_t\beta_t)) - ((\mathbb{P}\nabla \cdot (\beta_t \otimes \beta_t) + K'[\omega_t], v_t - \omega_t))$$

and the proof of Theorem 6.1 is now complete.

References

[1] L. Ambrosio, Transport equation and Cauchy problem for BV vector fields. *Invent. Math.* **158** (2004) 227–260.

[2] L. Ambrosio, N. Gigli and G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.* **195** (2014) 289–391.

[3] S. Angenent, S. Haker, A. Tannenbaum and R. Kikinis, On area preserving mappings of minimal distortion, System Theory: Modeling, Analysis and Control, Kluwer (2000) 275–286.

[4] V.I. Arnold and B. Khesin, Topological methods in hydrodynamics. Vol. 125 of *Appl. Math. Sci.* Springer-Verlag (1998).

[5] M. Balinski, A competitive (dual) simplex method for the assignment problem. *Math. Program.* **34** (1986) 125–141.

[6] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge−Kantorovich mass transfer problem. *Numer. Math.* **84** (2000) 375–393.

[7] T.B. Benjamin, The alliance of practical and analytical insight into the nonlinear problems of fluid mechanics. Vol. 53 of *Lect. Notes Math.* Springer-Verlag (1976) 8–29.

[8] G.R. Burton, Rearrangements of functions, maximization of convex functionals and vortex rings. *Math. Ann.* **276** (1987) 225–253

[9] Y. Brenier, Topology-preserving diffusion of divergence-free vector fields and magnetic relaxation. *Commun. Math. Phys.* **330** (2014) 757–770; *Linear Algebra Appl.* **146** (1991) 79–91.

[10] R.J. DiPerna and P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989) 511–547.

[11] F. Gay−Balmaz and D. Holm, Selective decay by Casimir dissipation in inviscid fluids. *Nonlinearity* **26** (2013) 495–524.

[12] D. Lesesvre, P. Pegon and F. Santambrogio, Optimal transportation with an oscillation-type cost: the one-dimensional case. *Set-Valued Var. Anal.* **21** (2013) 541–556.

[13] P.-L. Lions, Mathematical topics in fluid mechanics. 1. Incompressible models. Vol. 3 of *Oxford Lect. Series Math. Appl.* (1996).

[14] J. Louet and F. Santambrogio, A sharp inequality for transport maps in W1,p(R) *via* approximation. *Appl. Math. Lett.* **25** (2012) 648–653.

[15] C. Marchioro and M. Pulvirenti, Mathematical theory of incompressible nonviscous fluids. Springer-Verlag (1994).

[16] F. Mémoli, Some properties of Gromov-Hausdorff distances. *Discrete Comput. Geom.* **48** (2012) 416–440.

[17] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26** (2001) 101–174.

[18] K.T. Sturm, The space of spaces: curvature bounds and gradient flows on the space of metric measure spaces. Preprint `arXiv:1208.0434`.

[19] C. Villani, Optimal Transport, Old and New. Springer-Verlag (2009).