Geodesics in the space of measure-preserving maps and plans *

Luigi Ambrosio †
Scuola Normale Superiore
Pisa, Italy

Alessio Figalli ‡
Centre de Mathématiques Laurent Schwartz - Ecole Polytechnique
Palaiseau, France

Abstract

We study Brenier's variational models for incompressible Euler equations. These models give rise to a relaxation of the Arnold distance in the space of measure-preserving maps and, more generally, measure-preserving plans. We analyze the properties of the relaxed distance, we show a close link between the Lagrangian and the Eulerian model, and we derive necessary and sufficient optimality conditions for minimizers. These conditions take into account a modified Lagrangian induced by the pressure field. Moreover, adapting some ideas of Shnirelman, we show that, even for non-deterministic final conditions, generalized flows can be approximated in energy by flows associated to measure-preserving maps.

1 Introduction

The velocity of an incompressible fluid moving inside a region D is mathematically described by a time-dependent and divergence-free vector field $\boldsymbol{u}(t,x)$ which is parallel to the boundary ∂D. The Euler equations for incompressible fluids describes the evolution of such velocity field $\boldsymbol{u}$ in terms of the pressure field p:

$$\begin{cases} \partial_t \boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\nabla p & \text{in } [0,T]\times D, \\ \operatorname{div}\boldsymbol{u} = 0 & \text{in } [0,T]\times D, \\ \boldsymbol{u}\cdot n = 0 & \text{on } [0,T]\times \partial D. \end{cases} \tag{1.1}$$

Let us assume that $\boldsymbol{u}$ is smooth, so that it produces a unique flow g, given by

$$\begin{cases} \dot g(t,a) = \boldsymbol{u}(t,g(t,a)), \\ g(0,a) = a. \end{cases}$$

*This work was partially supported by the MIUR COFIN04 grant
†l.ambrosio@sns.it
‡figalli@math.polytechnique.fr

By the incompressibility condition, we get that at each time t the map $g(t,\cdot) : D \to D$ is a measure-preserving diffeomorphism of D, that is

$$g(t,\cdot)_\# \mu_D = \mu_D,$$

(here and in the sequel $f_\# \mu$ is the push-forward of a measure μ through a map f, and μ_D is the volume measure of the manifold D). Writing Euler equations in terms of g, we get

$$\begin{cases} \ddot{g}(t,a) = -\nabla p\,(t, g(t,a)) & (t,a) \in [0,T] \times D, \\ g(0,a) = a & a \in D, \\ g(t,\cdot) \in \mathrm{SDiff}(D) & t \in [0,T]. \end{cases} \tag{1.2}$$

Viewing the space SDiff(D) of measure-preserving diffeomorphisms of D as an infinite-dimensional manifold with the metric inherited from the embedding in L^2, and with tangent space made by the divergence-free vector fields, Arnold interpreted the equation above, and therefore (1.1), as a *geodesic* equation on SDiff(D) [6]. According to this intepretation, one can look for solutions of (1.2) by minimizing

$$T \int_0^T \int_D \frac{1}{2} |\dot{g}(t,x)|^2 \, d\mu_D(x) \, dt \tag{1.3}$$

among all paths $g(t,\cdot) : [0,T] \to \mathrm{SDiff}(D)$ with $g(0,\cdot) = f$ and $g(T,\cdot) = h$ prescribed (typically, by right invariance, f is taken as the identity map $\boldsymbol{i}$), and the pressure field arises as a Lagrange multiplier from the incompressibility constraint (the factor T in front of the integral is just to make the functional scale invariant in time). We shall denote by $\delta(f,h)$ the Arnold distance in SDiff(D), whose square is defined by the above-mentioned variational problem in the time interval $[0,1]$.

Although in the traditional approach to (1.1) the initial velocity is prescribed, while in the minimization of (1.3) is not, this variational problem has an independent interest and leads to deep mathematical questions, namely existence of relaxed solutions, gap phenomena and necessary and sufficient optimality conditions, that are investigated in this paper. We also remark that *no* existence result of distributional solutions of (1.1) is known when $d > 2$ (the case $d = 2$ is different, thanks to the vorticity formulation of (1.1)), see [21], [17] for a discussion on this topic and other concepts of weak solutions to (1.1).

On the positive side, Ebin and Marsden proved in [20] that, when D is a smooth compact manifold with no boundary, the minimization of (1.3) leads to a unique solution, corresponding also to a solution to Euler equations, if f and h are sufficienly close in a suitable Sobolev norm.

On the negative side, Shnirelman proved in [23], [24] that when $d \geq 3$ the infimum is not attained in general, and that when $d = 2$ there exists $h \in \mathrm{SDiff}(D)$ which cannot be connected to $\boldsymbol{i}$ by a path with finite action. These "negative" results motivate the study of relaxed versions of Arnold's problem.

The first relaxed version of Arnold's minimization problem was introduced by Brenier in [12]: he considered probability measures $\boldsymbol{\eta}$ in $\Omega(D)$, the space of continuous paths $\omega : [0,T] \to D$, and minimized the energy

$$\mathscr{A}_T(\boldsymbol{\eta}) := T \int_{\Omega(D)} \int_0^T \frac{1}{2} |\dot{\omega}(\tau)|^2 \, d\tau \, d\boldsymbol{\eta}(\omega),$$

with the constraints

$$(e_0, e_T)_\# \boldsymbol{\eta} = (\boldsymbol{i}, h)_\# \mu_D, \qquad (e_t)_\# \boldsymbol{\eta} = \mu_D \quad \forall t \in [0, T] \tag{1.4}$$

(here and in the sequel $e_t(\omega) := \omega(t)$ are the evaluation maps at time t). According to Brenier, we shall call these $\boldsymbol{\eta}$ *generalized incompressible flows* in $[0, T]$ between $\boldsymbol{i}$ and h. Obviously any sufficiently regular path $g(t, \cdot) : [0, 1] \to S(D)$ induces a generalized incompressible flow $\boldsymbol{\eta} = (\Phi_g)_\# \mu_D$, where $\Phi_g : D \to \Omega(D)$ is given by $\Phi_g(x) = g(\cdot, x)$, but the converse is far from being true: the main difference between classical and generalized flows consists in the fact that fluid paths starting from different points are allowed to cross at a later time, and fluid paths starting from the same point are allowed to split at a later time. This approach is by now quite common, see for instance [2] (DiPerna-Lions theory), [11] (branched optimal transportation), [22], [27].

Brenier's formulation makes sense not only if $h \in \mathrm{SDiff}(D)$, but also when $h \in S(D)$, where $S(D)$ is the space of measure-preserving maps $h : D \to D$, not necessarily invertible or smooth. In the case $D = [0, 1]^d$, existence of admissible paths with finite action connecting $\boldsymbol{i}$ to any $h \in S(D)$ was proved in [12], together with the existence of paths with minimal action. Furthermore, a consistency result was proved: smooth solutions to (1.1) are optimal even in the larger class of the generalized incompressible flows, provided the pressure field p satisfies

$$T^2 \sup_{t \in [0,T]} \sup_{x \in D} |\nabla_x^2 p(t, x)| \leq \pi^2, \tag{1.5}$$

and are the unique ones if the inequality is strict. When $\boldsymbol{\eta} = (\Phi_g)_\# \mu_D$ we can recover $g(t, \cdot)$ from $\boldsymbol{\eta}$ using the identity

$$(e_0, e_t)_\# \boldsymbol{\eta} = (\boldsymbol{i}, g(t, \cdot))_\# \mu_D, \qquad t \in [0, T].$$

Brenier found in [12] examples of action-minimizing paths $\boldsymbol{\eta}$ (for instance in the unit ball of $\mathbb{R}^2$, between $\boldsymbol{i}$ and $-\boldsymbol{i}$) where no such representation is possible. The same examples show that the upper bound (1.5) is sharp. Notice however that $(e_0, e_t)_\# \boldsymbol{\eta}$ is a measure-preserving plan, i.e. a probability measure in $D \times D$ having both marginals equal to μ_D. Denoting by $\Gamma(D)$ the space of measure-preserving plans, it is therefore natural to consider $t \mapsto (e_0, e_t)_\# \boldsymbol{\eta}$ as a "minimizing geodesic" between $\boldsymbol{i}$ and h in the larger space of measure-preserving plans. Then, to be consistent, one has to extend Brenier's minimization problem considering paths connecting γ, $\eta \in \Gamma(D)$. We define this extension, that reveals to be useful also to connect this model to the Eulerian-Lagrangian one in [16], and to obtain necessary and sufficient optimality conditions even when only "deterministic" data $\boldsymbol{i}$ and h are considered (because, as we said, the path might be non-deterministic in between). In this presentation of our results, however, to simplify the matter as much as possible, we shall consider the case of paths $\boldsymbol{\eta}$ between $\boldsymbol{i}$ and $h \in S(D)$ only.

In Section 5 we study the relation between the relaxation δ_* of the Arnold distance, defined by

$$\delta_*(h) := \inf \left\{ \liminf_{n \to \infty} \delta(\boldsymbol{i}, h_n) : \ h_n \in \mathrm{SDiff}(D), \ \int_D |h_n - h|^2 \, d\mu_D \to 0 \right\},$$

and the distance $\overline{\delta}(\boldsymbol{i},h)$ arising from the minimization of the Lagrangian model. It is not hard to show that $\overline{\delta}(\boldsymbol{i},h) \le \delta_*(h)$, and a natural question is whether equality holds, or a gap phenomenon occurs. In the case $D=[0,1]^d$ with $d>2$, an important step forward was obtained by Shnirelman in [24], who proved that equality holds when $h \in \mathrm{SDiff}(D)$; Shnirelman's construction provides an approximation (with convergence of the action) of generalized flows connecting $\boldsymbol{i}$ to h by smooth flows still connecting $\boldsymbol{i}$ to h. The main result of this section is the proof that no gap phenomenon occurs, still in the case $D=[0,1]^d$ with $d>2$, even when non-deterministic final data (i.e. measure-preserving plans) are considered. The proof of this fact is based on an auxiliary approximation result, Theorem 5.3, valid in any number of dimensions, which we believe of independent interest: it allows to approximate, with convergence of the action, any generalized flow $\boldsymbol{\eta}$ in $[0,1]^d$ by $W^{1,2}$ flows (in time) induced by measure-preserving maps $g(t,\cdot)$. This fact shows that the "negative" result of Shnirelman on the existence in dimension 2 of non-attainable diffeomorphisms is due to the regularity assumption on the path, and it is false if one allows for paths in the larger space $S(D)$. The proof of Theorem 5.3 uses some key ideas from [24] (in particular the combination of law of large numbers and smoothing of discrete families of trajectories), and some ideas coming from the theory of optimal transportation.

Minimizing generalized paths $\boldsymbol{\eta}$ are not unique in general, as shown in [12]; however, Brenier proved in [14] that the gradient of the pressure field p, identified by the distributional relation

$$\nabla p(t,x) = -\partial_t \overline{\boldsymbol{v}}_t(x) - \mathrm{div}\left(\overline{\boldsymbol{v}\otimes\boldsymbol{v}}_t(x)\right), \tag{1.6}$$

is indeed unique. Here $\overline{\boldsymbol{v}}_t(x)$ is the "effective velocity", defined by $(e_t)_\#(\dot\omega(t)\boldsymbol{\eta}) = \overline{\boldsymbol{v}}_t\mu_D$, and $\overline{\boldsymbol{v}\otimes\boldsymbol{v}}_t$ is the quadratic effective velocity, defined by $(e_t)_\#(\dot\omega(t)\otimes\dot\omega(t)\boldsymbol{\eta}) = \overline{\boldsymbol{v}\otimes\boldsymbol{v}}_t\mu_D$. The proof of this fact is based on the so-called dual least action principle: if $\boldsymbol{\eta}$ is optimal, we have

$$\mathscr{A}_T(\boldsymbol{\nu}) \ge \mathscr{A}_T(\boldsymbol{\eta}) + \langle p, \rho^{\boldsymbol{\nu}} - 1\rangle \tag{1.7}$$

for any measure $\boldsymbol{\nu}$ in $\Omega(D)$ such that $(e_0,e_T)_\#\boldsymbol{\nu} = (\boldsymbol{i},h)_\#\mu_D$ and $\|\rho^{\boldsymbol{\nu}} - 1\|_{C^1} \le 1/2$. Here $\rho^{\boldsymbol{\nu}}$ is the (absolutely continuous) density produced by the flow $\boldsymbol{\nu}$, defined by $\rho^{\boldsymbol{\nu}}(t,\cdot)\mu_D = (e_t)_\#\boldsymbol{\nu}$. In this way, the incompressibility constraint can be slightly relaxed and one can work with the augmented functional (still minimized by $\boldsymbol{\eta}$)

$$\boldsymbol{\nu} \mapsto \mathscr{A}_T(\boldsymbol{\nu}) - \langle p, \rho^{\boldsymbol{\nu}} - 1\rangle,$$

whose first variation leads to (1.6).

In Theorem 6.2, still using the key Proposition 2.1 from [14], we provide a simpler proof and a new interpretation of the dual least action principle.

A few years later, Brenier introduced in [16] a new relaxed version of Arnold's problem of a mixed Eulerian-Lagrangian nature: the idea is to add to the Eulerian variable x a Lagrangian one a representing, at least when $f=\boldsymbol{i}$, the initial position of the particle; then, one minimizes a functional of the Eulerian variables (density and velocity), depending also on a. Brenier's motivation for looking at the new model was that this formalism allows to show much stronger regularity results for the pressure field, namely $\partial_{x_i}p$ are locally finite measures in $(0,T)\times D$. In Section 3.3 we describe in detail this new model and, in Section 4, we show that the two models are basically equivalent. This result will be used by us to transfer the regularity informations on the pressure field up to the Lagrangian model, thus obtaining the validity of (1.7) for a much larger class of generalized flows $\boldsymbol{\nu}$, that we call flows with bounded compression. The proof of the equivalence follows by a general principle (Theorem 2.4, borrowed from [3]) that allows to move from an Eulerian to a Lagrangian description, lifting solutions to the continuity equation to measures in the space of continuous maps.

In the final section of our paper we look for necessary and sufficient optimality conditions for the geodesic problem. These conditions require that the pressure field p is a function and not only a distribution: this technical result is achieved in [5], where, by carefully analyzing and improving Brenier's difference-quotient argument, we show that $\partial_{x_i} p \in L^2_{\rm loc}\big((0,T);\mathcal{M}_{\rm loc}(D)\big)$ (this implies, by Sobolev embedding, $p \in L^2_{\rm loc}\big((0,T);L^{d/(d-1)}_{\rm loc}(D)\big)$).

In this final section, although we do not see a serious obstruction to the extension of our results to a more general framework, we consider the case of the flat torus $\mathbb{T}^d$ only, and we shall denote by $\mu_{\mathbb{T}}$ the canonical measure on the flat torus. We observe that in this case $p \in L^2_{\rm loc}\big((0,T);L^{d/(d-1)}(\mathbb{T}^d)\big)$ and so, taking into account that the pressure field in (1.7) is uniquely determined up to additive time-dependent constants, we may assume that $\int_{\mathbb{T}^d} p(t,\cdot) d\mu_{\mathbb{T}} = 0$ for almost all $t \in (0,T)$.
The first elementary remark is that any integrable function q in $(0,T)\times\mathbb{T}^d$ with $\int_{\mathbb{T}^d} q(t,\cdot)\,d\mu_{\mathbb{T}} = 0$ for almost all $t \in (0,T)$ provides us with a null-lagrangian for the geodesic problem, as the incompressibility constraint gives

$$\int_{\Omega(\mathbb{T}^d)} \int_0^T q(t,\omega(t))\,dt\,d\boldsymbol{\nu}(\omega) = \int_0^T \int_{\mathbb{T}^d} q(t,x)\,d\mu_{\mathbb{T}}(x)\,dt = 0$$

for any generalized incompressible flow $\boldsymbol{\nu}$. Taking also the constraint $(e_0,e_T)_\#\boldsymbol{\nu} = (\boldsymbol{i},h)_\#\mu$ into account, we get

$$\mathscr{A}_T(\boldsymbol{\nu}) = T\int_{\Omega(\mathbb{T}^d)} \left(\int_0^T \frac{1}{2}|\dot\omega(t)|^2 - q(t,\omega)\,dt\right) d\boldsymbol{\nu}(\omega) \ge \int_{\mathbb{T}^d} c_q^T(x,h(x))\,d\mu_{\mathbb{T}}(x),$$

where $c_q^T(x,y)$ is the minimal cost associated with the Lagrangian $T\int_0^T \frac{1}{2}|\dot\omega(t)|^2 - q(t,\omega)\,dt$. Since this lower bound depends only on h, we obtain that any $\boldsymbol{\eta}$ satisfying (1.4) and concentrated on c_q-minimal paths, for some $q \in L^1$, is optimal, and $\overline{\delta}^2(\boldsymbol{i},h) = \int c_q^T(\boldsymbol{i},h)\,d\mu_{\mathbb{T}}$. This is basically the argument used by Brenier in [12] to show the minimality of smooth solutions to (1.1), under assumption (1.5): indeed, this condition guarantees that solutions of $\ddot\omega(t) = -\nabla p(t,\omega)$ (i.e. stationary paths for the Lagrangian, with $q = p$) are also minimal.

We are able to show that basically this condition is *necessary and sufficient* for optimality if the pressure field is globally integrable (see Theorem 6.12). However, since no global in time regularity result for the pressure field is presently known, we have also been looking for necessary and sufficient optimality conditions that don't require the global integrability of the pressure field. Using the regularity $p \in L^1_{\rm loc}\left((0,T);L^r(D)\right)$ for some $r > 1$, guaranteed in the case $D = \mathbb{T}^d$ with $r = d/(d-1)$ by [5], we show in Theorem 6.8 that any optimal $\boldsymbol{\eta}$ is concentrated on *locally* *minimizing* paths for the Lagrangian

$$\mathcal{L}_p(\omega) := \int \frac{1}{2}|\dot\omega(t)|^2 - p(t,\omega)\,dt \tag{1.8}$$

Since we are going to integrate p along curves, this statement is not invariant under modifications of p in negligible sets, and the choice of a *specific* representative $\bar p(t,x) := \liminf_{\varepsilon\downarrow 0} p(t,\cdot) * \phi_\varepsilon(x)$ in the Lebesgue equivalence class is needed. Moreover, the necessity of pointwise uniform estimates on p_ε requires the integrability of $Mp(t,x)$, the maximal function of $p(t,\cdot)$ at x (see (6.11)).

In addition, we identify a second necessary (and more hidden) optimality condition. In order to state it, let us consider an interval $[s,t]\subset(0,T)$ and the cost function

$$c^{s,t}_p(x,y) := \inf\left\{\int_s^t \frac{1}{2}|\dot\omega(\tau)|^2 - p(\tau,\omega)\,d\tau:\ \omega(s)=x,\, \omega(t)=y,\, Mp(\tau,\omega)\in L^1(s,t)\right\}. \tag{1.9}$$

(the assumption $Mp(\tau,\omega)\in L^1(s,t)$ is forced by technical reasons). Recall that, according to the theory of optimal transportation, a probability measure λ in $\mathbb{T}^d\times\mathbb{T}^d$ is said to be c-optimal if

$$\int_{\mathbb{T}^d\times\mathbb{T}^d} c(x,y)\,d\lambda' \geq \int_{\mathbb{T}^d\times\mathbb{T}^d} c(x,y)\,d\lambda$$

for any probability measure λ' having the same marginals μ_1, μ_2 of λ. We shall also denote $W_c(\mu_1,\mu_2)$ the minimal value, i.e. $\int_{\mathbb{T}^d\times\mathbb{T}^d} c\,d\lambda$, with λ c-optimal. Now, let $\boldsymbol\eta$ be an optimal generalized incompressible flow between $\boldsymbol i$ and h; according to the disintegration theorem, we can represent $\boldsymbol\eta = \int \boldsymbol\eta_a\,d\mu_D(a)$, with $\boldsymbol\eta_a$ concentrated on curves starting at a (and ending, since our final conditions is deterministic, at $h(a)$), and consider the plans $\lambda^{s,t}_a = (e_s,e_t)_\#\boldsymbol\eta_a$. We show that

$$\text{for all } [s,t]\subset(0,T), \quad \lambda^{s,t}_a \text{ is } c^{s,t}_p\text{-optimal for } \mu_{\mathbb{T}}\text{-a.e. } a\in\mathbb{T}^d. \tag{1.10}$$

Roughly speaking, this condition tells us that one has not only to move mass from x to y achieving $c^{s,t}_p$, but also to optimize the distribution of mass between time s and time t. In the "deterministic" case when either $(e_0,e_s)_\#\boldsymbol\eta$ or $(e_0,e_t)_\#\boldsymbol\eta$ are induced by a transport map g, the plan $\lambda^{s,t}_a$ has $\delta_{g(a)}$ either as first or as second marginal, and therefore it is uniquely determined by its marginals (it is indeed the product of them). This is the reason why condition (1.10) does not show up in the deterministic case.

Finally, we show in Theorem 6.12 that the two conditions are also sufficient, even on general manifolds D: if, for some $r>1$ and $q\in L^1_{\rm loc}((0,T);L^r(D))$, a generalized incompressible flow $\boldsymbol\eta$ concentrated on locally minimizing curves for the Lagrangian $\mathcal{L}_q$ satisfies

$$\text{for all } [s,t]\subset(0,T), \quad \lambda^{s,t}_a \text{ is } c^{s,t}_q\text{-optimal for } \mu_D\text{-a.e. } a\in D,$$

then $\boldsymbol\eta$ is optimal in $[0,T]$, and q is the pressure field.

These results show a somehow unexpected connection between the variational theory of incompressible flows and the theory developed by Bernard-Buffoni [9] of measures in the space of action-minimizing curves; in this framework one can fit Mather's theory as well as optimal transportation problems on manifolds, with a geometric cost. In our case the only difference is that the Lagrangian is possibly nonsmoooth (but hopefully not so bad), and not given *a priori*, but generated by the problem itself. Our approach also yields (see Corollary 6.13) a new variational characterization of the pressure field, as a maximizer of the family of functionals (for $[s,t] \subset (0,T)$)

$$q \mapsto \int_{\mathbb{T}^d} W_{c_q^{s,t}}(\eta_a^s, \gamma_a^t)\, d\mu_{\mathbb{T}}(a), \qquad Mq \in L^1\left([s,t] \times \mathbb{T}^d\right),$$

where η_a^s, γ_a^t are the marginals of $\lambda_a^{s,t}$.

Acknowledgement. We warmly thank Yann Brenier for the many discussions we had together, and for the constant support we had from him. We also thank an anonymous referee, who read very carefully this paper, for his comments.

2 Notation and preliminary results

Measure-theoretic notation. We start by recalling some basic facts in Measure Theory. Let X, Y be *Polish* spaces, i.e. topological spaces whose topology is induced by a complete and separable distance. We endow a Polish space X with the corresponding Borel σ-algebra and denote by $\mathscr{P}(X)$ (resp. $\mathscr{M}_+(X)$, $\mathscr{M}(X)$) the family of Borel probability (resp. nonnegative and finite, real and with finite total variation) measures in X. For $A \subset X$ and $\mu \in \mathscr{M}(X)$ the *restriction* $\mu \llcorner A$ of μ to A is defined by $\mu \llcorner A(B) := \mu(A \cap B)$. We will denote by $\boldsymbol{i} : X \to X$ the identity map.

Definition 2.1 (Push-forward) *Let $\mu \in \mathscr{M}(X)$ and let $f : X \to Y$ be a Borel map. The push-forward $f_\#\mu$ is the measure in Y defined by $f_\#\mu(B) = \mu(f^{-1}(B))$ for any Borel set $B \subset Y$. The definition obviously extends, componentwise, to vector-valued measures.*

It is easy to check that $f_\#\mu$ has finite total variation as well, and that $|f_\#\mu| \leq f_\#|\mu|$. An elementary approximation by simple functions shows the change of variable formula

$$\int_Y g\, df_\#\mu = \int_X g \circ f\, d\mu \tag{2.1}$$

for any bounded Borel function (or even either nonnegative or nonpositive, and $\overline{\mathbb{R}}$-valued, in the case $\mu \in \mathscr{M}_+(X)$) $g : Y \to \mathbb{R}$.

Definition 2.2 (Narrow convergence and compactness) *Narrow (sequential) convergence in $\mathscr{P}(X)$ is the convergence induced by the duality with $C_b(X)$, the space of continuous and bounded functions in X. By Prokhorov theorem, a family $\mathscr{F}$ in $\mathscr{P}(X)$ is sequentially relatively compact with respect to the narrow convergence if and only if it is tight, i.e. for any $\varepsilon > 0$ there exists a compact set $K \subset X$ such that $\mu(X \setminus K) < \varepsilon$ for any $\mu \in \mathscr{F}$.*

In this paper we use only the "easy" implication in Prokhorov theorem, namely that any tight family is sequentially relatively compact. It is immediate to check that a sufficient condition for tightness of a family $\mathscr{F}$ of probability measures is the existence of a *coercive* functional $\Psi : X \to [0, +\infty]$ (i.e. a functional such that its sublevel sets $\{\Psi \le t\}$, $t \in \mathbb{R}^+$, are relatively compact in X) such that

$$\int_X \Psi(x) d\mu(x) \le 1 \qquad \forall \mu \in \mathscr{F}.$$

Lemma 2.3 ([4], Lemma 2.4) *Let $\mu \in \mathscr{P}(X)$ and $\boldsymbol{u} \in L^2(X; \mathbb{R}^m)$. Then, for any Borel map $f : X \to Y$, $f_\#(\boldsymbol{u}\mu) \ll f_\#\mu$ and its density $\boldsymbol{v}$ with respect to $f_\#\mu$ satisfies*

$$\int_Y |\boldsymbol{v}|^2 \, df_\#\mu \le \int_X |\boldsymbol{u}|^2 \, d\mu.$$

Furthermore, equality holds if and only if $\boldsymbol{u} = \boldsymbol{v} \circ f$ μ-a.e. in X.

Given $\mu \in \mathscr{M}_+(X \times Y)$, we shall denote by $\mu_x \otimes \lambda$ its *disintegration* via the projection map $\pi(x, y) = x$: here $\lambda = \pi_\#\mu \in \mathscr{M}_+(X)$, and $x \mapsto \mu_x \in \mathscr{P}(Y)$ is a Borel map (i.e. $x \mapsto \mu_x(A)$ is Borel for all Borel sets $A \subset Y$) characterized, up to λ-negligible sets, by

$$\int_{X \times Y} f(x, y) \, d\mu(x, y) = \int_X \left(\int_Y f(x, y) \, d\mu_x(y) \right) d\lambda(x) \tag{2.2}$$

for all nonnegative Borel map f. Conversely, any λ and any Borel map $x \mapsto \mu_x \in \mathscr{P}(Y)$ induce a probability measure μ in $X \times Y$ via (2.2).

Function spaces. We shall denote by $\Omega(D)$ the space $C([0, T]; D)$, and by $\omega : [0, T] \to D$ its typical element. The evaluation maps at time t, $\omega \mapsto \omega(t)$, will be denoted by e_t.

If D is a smooth, compact Riemannian manifold without boundary (typically the d-dimensional flat torus $\mathbb{T}^d$), we shall denote μ_D its volume measure, and by d_D its Riemannian distance, normalizing the Riemannian metric so that μ_D is a probability measure. Although it does not fit exactly in this framework, we occasionally consider also the case $D = [0, 1]^d$, because many results have already been obtained in this particular case.

We shall often consider measures $\boldsymbol{\eta} \in \mathscr{M}_+(\Omega(D))$ such that $(e_t)_\#\boldsymbol{\eta} \ll \mu_D$; in this case we shall denote by $\rho^{\boldsymbol{\eta}} : [0, T] \times D \to [0, +\infty]$ the density, characterized by

$$\rho^{\boldsymbol{\eta}}(t, \cdot)\mu_D := (e_t)_\#\boldsymbol{\eta}, \qquad t \in [0, T].$$

We denote by SDiff(D) the measure-preserving diffeomorphisms of D, and by $S(D)$ the measure-preserving maps in D:

$$S(D) := \{g : D \to D : \; g_\#\mu_D = \mu_D\}. \tag{2.3}$$

We also set

$$S^{\mathfrak{i}}(D) := \{g \in S(D) : g \text{ is } \mu_D\text{-essentially injective}\}. \tag{2.4}$$

For any $g \in S^{\mathfrak{i}}(D)$ the inverse g^{-1} is well defined up to μ_D-negligible sets, μ_D-measurable, and $g^{-1} \circ g = \boldsymbol{i} = g \circ g^{-1}$ μ_D-a.e. in D. In particular, if $g \in S^{\mathfrak{i}}(D)$, $g^{-1} \in S^{\mathfrak{i}}(D)$.

We shall also denote by $\Gamma(D)$ the family of measure-preserving plans, i.e. the probability measures in $D \times D$ whose first and second marginal are μ_D:

$$\Gamma(D) := \{\gamma \in \mathscr{P}(D \times D) : \ (\pi_1)_\# \gamma = \mu_D, \ (\pi_2)_\# \gamma = \mu_D\} \tag{2.5}$$

(here π_1, π_2 are the canonical coordinate projections).

Recall that $\mathrm{SDiff}(D) \subset S^i(D) \subset S(D)$ and that any element $g \in S(D)$ canonically induces a measure preserving plan γ_g, defined by

$$\gamma_g := (\boldsymbol{i} \times g)_\# \mu_D.$$

Furthermore, this correspondence is continuous, as long as convergence in $L^2(\mu)$ of the maps g and narrow convergence of the plans are considered (see for instance Lemma 2.3 in [4]). Moreover

$$\overline{\{\gamma_g : \ g \in S^i(D)\}}^{\text{narrow}} = \Gamma(D), \tag{2.6}$$

$$\overline{\mathrm{SDiff}(D)}^{L^2(\mu_D)} = S(D) \qquad \text{if } D = [0,1]^d, \text{ with } d \geq 2 \tag{2.7}$$

(the first result is standard, see for example the explicit construction in [18, Theorem 1.4 (i)] in the case $D = [0,1]^d$, while the second one is proved in [18, Corollary 1.5])

The continuity equation. In the sequel we shall often consider weak solutions $\mu_t \in \mathscr{P}(D)$ of the continuity equation

$$\partial_t \mu_t + \mathrm{div}(\boldsymbol{v}_t \mu_t) = 0, \tag{2.8}$$

where $t \mapsto \mu_t$ is narrowly continuous (this is not restrictive, see for instance Lemma 8.1.2 of [3]) and $\boldsymbol{v}_t(x)$ is a suitable velocity field with $\|\boldsymbol{v}_t\|_{L^2(\mu_t)} \in L^1(0,T)$ (formally, $\boldsymbol{v}_t$ is a section of the tangent bundle and $|\boldsymbol{v}_t|$ is computed according to the Riemannian metric). The equation is understood in a weak (distributional) sense, by requiring that

$$\frac{d}{dt} \int_D \phi(t,x)\, d\mu_t(x) = \int_D \partial_t \phi + \langle \nabla \phi, \boldsymbol{v}_t \rangle \, d\mu_t \qquad \text{in } \mathcal{D}'(0,T)$$

for any $\phi \in C^1\left((0,T) \times D\right)$ with bounded first derivatives and support contained in $J \times D$, with $J \Subset (0,T)$. In the case when $D \subset \mathbb{R}^d$ is compact, we shall consider functions $\phi \in C^1\left((0,T) \times \mathbb{R}^d\right)$, again with support contained in $J \times \mathbb{R}^d$, with $J \Subset (0,T)$.

The following general principle allows to lift solutions of the continuity equation to measures in the space of continuous paths.

Theorem 2.4 (Superposition principle) *Assume that either D is a compact subset of $\mathbb{R}^d$, or D is a smooth compact Riemannian manifold without boundary, and let $\mu_t : [0,T] \to \mathscr{P}(D)$ be a narrowly continuous solution of the continuity equation* (2.8) *for a suitable velocity field $\boldsymbol{v}(t,x) = \boldsymbol{v}_t(x)$ satisfying $\|\boldsymbol{v}_t\|^2_{L^2(\mu_t)} \in L^1(0,T)$. Then there exists $\boldsymbol{\eta} \in \mathscr{P}(\Omega(D))$ such that*

(i) $\mu_t = (e_t)_\# \boldsymbol{\eta}$ for all $t \in [0,T]$;

(ii) the following energy inequality holds:

$$\int_{\Omega(D)} \int_0^T |\dot{\omega}(t)|^2 \, dt \, d\boldsymbol{\eta}(\omega) \le \int_0^T \int_D |\boldsymbol{v}_t|^2 \, d\mu_t \, dt.$$

Proof. In the case when $D = \mathbb{R}^d$ (and therefore also when $D \subset \mathbb{R}^d$ is closed) this result is proved in Theorem 8.2.1 of [3] (see also [7], [25], [10] for related results). In the case when D is a smooth, compact Riemannian manifold we recover the same result thanks to an isometric embedding in $\mathbb{R}^m$, for m large enough. □

3 Variational models for generalized geodesics

3.1 Arnold's least action problem

Let f, $h \in \mathrm{SDiff}(D)$ be given. Following Arnold [6], we define $\delta^2(f,h)$ by minimizing the action

$$\mathscr{A}_T(g) := T \int_0^T \int_D \frac{1}{2} |\dot{g}(t,x)|^2 \, d\mu_D(x) \, dt,$$

among all smooth curves

$$[0,T] \ni t \mapsto g(t,\cdot) \in \mathrm{SDiff}(D)$$

connecting f to h. By time rescaling, δ is independent of T. Since right composition with a given element $g \in \mathrm{SDiff}(D)$ does not change the action (as it amounts just to a relabelling of the initial position with g), the distance δ is right invariant, so it will be often useful to assume, in the minimization problem, that f is the identity map.

The action $\mathscr{A}_T$ can also be computed in terms of the velocity field $\boldsymbol{u}$, defined by $\boldsymbol{u}(t,x) = \dot{g}(t,y)|_{y=g^{-1}(t,x)}$, as

$$\mathscr{A}_T(\boldsymbol{u}) = T \int_0^T \int_D \frac{1}{2} |\boldsymbol{u}(t,x)|^2 \, d\mu_D(x) \, dt.$$

As we mentioned in the introduction, connections between this minimization problem and (1.1) were achieved first by Ebin and Marsden, and then by Brenier: in [12], [16] he proved that if $(\boldsymbol{u},p)$ is a smooth solution of the Euler equation in $[0,T] \times D$, with $D = [0,1]^d$, and the inequality in (1.5) is strict, then the flow $g(t,x)$ of $\boldsymbol{u}$ is the unique solution of Arnold's minimization problem with $f = \boldsymbol{i}$, $h = g(T,\cdot)$.

By integrating the inequality $d_D^2(h(x), f(x)) \le \int_0^1 |\dot{g}(t,x)|^2 \, dt$ one immediately obtains that $\|h - f\|_{L^2(D)} \le \sqrt{2}\delta(f,h)$; Shnirelman proved in [24] that in the case $D = [0,1]^d$ with $d \ge 3$ the Arnold distance is topologically equivalent to the L^2 distance: namely, there exist $C > 0$, $\alpha > 0$ such that

$$\delta(f,g) \le C \|f - g\|_{L^2(D)}^{\alpha} \qquad \forall f,\, g \in \mathrm{SDiff}(D). \tag{3.1}$$

Shnirelman also proved in [23] that when $d \ge 3$ the infimum is not attained in general and that, when $d = 2$, $\delta(\boldsymbol{i}, h)$ need not be finite (i.e., there exist $h \in \mathrm{SDiff}(D)$ which cannot be connected to $\boldsymbol{i}$ by a path with finite action).

3.2 Brenier's Lagrangian model and its extensions

In [12], Brenier proposed a relaxed version of the Arnold geodesic problem, and here we present more general versions of Brenier's relaxed problem, allowing first for final data in $\Gamma(D)$, and then for initial and final data in $\Gamma(D)$.

Let $\gamma \in \Gamma(D)$ be given; the class of admissible paths, called by Brenier *generalized incompressible flows*, is made by the probability measures $\boldsymbol{\eta}$ on $\Omega(D)$ such that

$$(e_t)_\# \boldsymbol{\eta} = \mu_D \qquad \forall t \in [0,T].$$

Then the action of an admissible $\boldsymbol{\eta}$ is defined as

$$\mathscr{A}_T(\boldsymbol{\eta}) := \int_{\Omega(D)} \mathscr{A}_T(\omega)\, d\boldsymbol{\eta}(\omega),$$

where

$$\mathscr{A}_T(\omega) := \begin{cases} T\int_0^T \frac{1}{2}|\dot\omega(t)|^2\,dt & \text{if } \omega \text{ is absolutely continuous in } [0,T] \\ +\infty & \text{otherwise,} \end{cases} \tag{3.2}$$

and $\overline{\delta}^2(\gamma_{\boldsymbol{i}},\gamma)$ is defined by minimizing $\mathscr{A}_T(\boldsymbol{\eta})$ among all generalized incompressible flows $\boldsymbol{\eta}$ connecting $\gamma_{\boldsymbol{i}}$ to γ, i.e. those satisfying

$$(e_0, e_T)_\# \boldsymbol{\eta} = \gamma. \tag{3.3}$$

Notice that it is not clear, in this purely Lagrangian formulation, how the relaxed distance $\overline{\delta}(\eta,\gamma)$ between two measure preserving plans might be defined, not even when η and γ are induced by maps g, h. Only when $g \in S^i(D)$ we might use the right invariance and define $\overline{\delta}(\gamma_g,\gamma_h) := \overline{\delta}(\gamma_{\boldsymbol{i}},\gamma_{h\circ g^{-1}})$.

These remarks led us to the following more general problem: let us denote

$$\tilde\Omega(D) := \Omega(D) \times D,$$

whose typical element will be denoted by (ω,a), and let us denote by $\pi_D : \tilde\Omega(D) \to D$ the canonical projection. We consider probability measures $\boldsymbol{\eta}$ in $\tilde\Omega(D)$ having μ_D as second marginal, i.e. $(\pi_D)_\# \boldsymbol{\eta} = \mu_D$; they can be canonically represented as $\boldsymbol{\eta}_a \otimes \mu_D$, where $\boldsymbol{\eta}_a \in \mathscr{P}(\Omega(D))$. The incompressibility constraint now becomes

$$\int_D (e_t)_\# \boldsymbol{\eta}_a \, d\mu_D(a) = \mu_D \qquad \forall t \in [0,T], \tag{3.4}$$

or equivalently $(e_t)_\# \boldsymbol{\eta} = \mu_D$ for all t, if we consider e_t as a map defined on $\tilde\Omega(D)$. Given initial and final data $\eta = \eta_a \otimes \mu_D$, $\gamma = \gamma_a \otimes \mu_D \in \Gamma(D)$, the constraint (3.3) now becomes

$$(e_0, \pi_D)_\# \boldsymbol{\eta} = \eta_a \otimes \mu_D, \qquad (e_T, \pi_D)_\# \boldsymbol{\eta} = \gamma_a \otimes \mu_D. \tag{3.5}$$

Equivalently, in terms of $\boldsymbol{\eta}_a$ we can write

$$(e_0)_\# \boldsymbol{\eta}_a = \eta_a, \qquad (e_T)_\# \boldsymbol{\eta}_a = \gamma_a. \tag{3.6}$$

Then, we define $\overline{\delta}^2(\eta,\gamma)$ by minimizing the action

$$\int_{\tilde{\Omega}(D)} \mathscr{A}_T(\omega)\, d\boldsymbol{\eta}(\omega,a)$$

among all generalized incompressible flows $\boldsymbol{\eta}$ (according to (3.4)) connecting η to γ (according to (3.5) or (3.6)). Notice that $\overline{\delta}^2$ is independent of T, because the action is scaling invariant; so we can use any interval $[a,b]$ in place of $[0,T]$ to define $\overline{\delta}$, and in this case we shall talk of generalized flow between η and γ in $[a,b]$ (this extension will play a role in Remark 3.2 below).

When $\eta_a = \delta_a$ (i.e. $\eta = \gamma_{\boldsymbol{i}}$), (3.6) tells us that almost all trajectories of $\boldsymbol{\eta}_a$ start from a: then $\int_D \boldsymbol{\eta}_a\, d\mu_D(a)$ provides us with a solution of Brenier's original model with the same action, connecting $\gamma_{\boldsymbol{i}}$ to γ. Conversely, any solution $\boldsymbol{\nu}$ of this model can be written as $\int_D \boldsymbol{\nu}_a\, d\mu_D$, with $\boldsymbol{\nu}_a$ concentrated on the curves starting at a, and $\boldsymbol{\nu}_a \otimes \mu_D$ provides us with an admissible path for our generalized problem, connecting $\gamma_{\boldsymbol{i}}$ to γ, with the same action.

Let us now analyze the properties of $(\Gamma(D),\overline{\delta})$; the fact that this is a metric space and even a *length space* (i.e. any two points can be joined by a geodesic with length equal to the distance) follows by the basic operations *reparameterization*, *restriction* and *concatenation* of generalized flows, that we are now going to describe.

Remark 3.1 (Repameterization) Let $\chi : [0,T] \to [0,T]$ be a C^1 map with $\dot\chi > 0$, $\chi(0) = 0$ and $\chi(T) = T$. Then, right composition of ω with χ induces a transformation $\boldsymbol{\eta} \mapsto \chi_\#\boldsymbol{\eta}$ between generalized incompressible flows that preserves the initial and final conditionl. As a consequence, if $\boldsymbol{\eta}$ is optimal the functional $\chi \mapsto \mathscr{A}_T(\chi_\#\boldsymbol{\eta})$ attains its minimum when $\chi(t) = t$. Changing variables we obtain

$$\mathscr{A}_T(\chi_\#\boldsymbol{\eta}) = T\int_0^T \dot\chi^2(t)\int_{\tilde{\Omega}(D)} \frac{1}{2}|\dot\omega|^2(\chi(t))\, d\boldsymbol{\eta}(\omega,a)\, dt = T\int_0^T \frac{1}{\dot g(s)}\int_{\tilde{\Omega}(D)} \frac{1}{2}|\dot\omega|^2(s)\, d\boldsymbol{\eta}(\omega,a)\, ds$$

with $g = \chi^{-1}$. Therefore, choosing $g(s) = s + \varepsilon\phi(s)$, with $\phi \in C^1_c(0,T)$, the first variation gives

$$\int_0^T \left(\int_{\tilde{\Omega}(D)} |\dot\omega|^2(s)\, d\boldsymbol{\eta}(\omega,a)\right)\dot\phi(s)\, ds = 0.$$

This proves that $s \mapsto \int_{\tilde{\Omega}(D)} |\dot\omega|^2(s)\, d\boldsymbol{\eta}(\omega,a)$ is equivalent to a constant. We shall call the square root of this quantity *speed* of $\boldsymbol{\eta}$.

Remark 3.2 (Restriction and concatenation) Let $[s,t] \subset [0,T]$ and let $r_{s,t} : C([0,T];D) \to C([s,t];D)$ be the restriction map. It is immediate to check that, for any generalized incompressible flow $\boldsymbol{\eta} = \boldsymbol{\eta}_a \otimes \mu_D$ in $[0,T]$ between η and γ, the measure $(r_{s,t})_\#\boldsymbol{\eta}$ is a generalized incompressible flow in $[s,t]$ between $\eta_s := (e_s)_\#\boldsymbol{\eta}_a \otimes \mu_D$ and $\gamma_t := (e_t)_\#\boldsymbol{\eta}_a \otimes \mu_D$, with action equal to

$$(t-s)\int_{\tilde{\Omega}(D)}\int_s^t \frac{1}{2}|\dot\omega(\tau)|^2\, d\tau\, d\boldsymbol{\eta}(\omega,a).$$

Let $s < l < t$ and let $\boldsymbol{\eta} = \boldsymbol{\mu}_a \otimes \mu_D$, $\boldsymbol{\nu} = \boldsymbol{\nu}_a \otimes \mu_D$ be generalized incompressible flows, respectively defined in $[s,l]$ and $[l,t]$, and joining η to γ and γ to θ. Then, writing $\gamma_a = (e_l)_\# \boldsymbol{\eta}_a = (e_l)_\# \boldsymbol{\nu}_a$, we can disintegrate both $\boldsymbol{\eta}_a$ and $\boldsymbol{\nu}_a$ with respect to γ_a to obtain

$$\boldsymbol{\eta}_a = \int_D \boldsymbol{\eta}_{a,x}\, d\gamma_a(x) \in \mathscr{P}(C([s,l];D)), \qquad \boldsymbol{\nu}_a = \int_D \boldsymbol{\nu}_{a,x}\, d\gamma_a(x) \in \mathscr{P}(C([l,t];D)),$$

with $\boldsymbol{\eta}_{a,x}$, $\boldsymbol{\nu}_{a,x}$ concentrated on the curves ω with $\omega(l) = x$. We can then consider the image $\boldsymbol{\lambda}_{x,a}$, via the concatenation of paths (from the product of $C([s,l];D)$ and $C([l,t];D)$ to $C([s,t];D)$), of the product measure $\boldsymbol{\eta}_{a,x} \times \boldsymbol{\nu}_{a,x}$ to obtain a probability measure in $C([s,t];D)$ concentrated on paths passing through x at time l. Eventually, setting

$$\boldsymbol{\lambda} = \int_{D\times D} \boldsymbol{\lambda}_{x,a}\, d(\gamma_a \otimes \mu_D)(x,a),$$

we obtain a generalized incompressible flow in $[s,t]$ joining η to θ with action given by

$$\frac{t-s}{l-s}\mathscr{A}_{[s,l]}(\boldsymbol{\eta}) + \frac{t-s}{t-l}\mathscr{A}_{[l,t]}(\boldsymbol{\nu}),$$

where $\mathscr{A}_{[s,l]}(\boldsymbol{\eta})$ is the action of $\boldsymbol{\eta}$ in $[s,l]$ and $\mathscr{A}_{[l,t]}(\boldsymbol{\nu})$ is the action of $\boldsymbol{\nu}$ in $[l,t]$ (strictly speaking, the action of their restrictions).

A simple consequence of the previous remarks is that $\overline{\delta}$ is a distance in $\Gamma(D)$ (it suffices to concatenate flows with unit speed); in addition, the restriction of an optimal incompressible flow $\boldsymbol{\eta} = \boldsymbol{\eta}_a \otimes \mu_D$ between $\eta_a \otimes \mu_D$ and $\gamma_a \otimes \mu_D$ to an interval $[s,t]$ is still an optimal incompressible flow in $[s,t]$ between the plans $(e_s)_\# \boldsymbol{\eta}_a \otimes \mu_D$ and $(e_t)_\# \boldsymbol{\eta}_a \otimes \mu_D$. This property will be useful in Section 6.

Another important property of $\overline{\delta}$ that will be useful in Section 6 is its lower semicontinuity with respect to the narrow convergence, that we are going to prove in the next theorem. Another non-trivial fact is the existence of at least one generalized incompressible flow with finite action. In [12, Section 4] Brenier proved the existence of such a flow in the case $D = \mathbb{T}^d$. Then in [24, Section 2], using a (non-injective) Lipschitz measure-preserving map from $\mathbb{T}^d$ to $[0,1]^d$, Shnirelman produced a flow with finite action also in this case (see also [16, Section 3]). In the next theorem we will show how to construct a flow with finite action in a compact subset D whenever flows with finite action can be built in D' and a possibly non-injective, Lipschitz and measure-preserving map $f : D' \to D$ exists.

Theorem 3.3 *Assume that $D \subset \mathbb{R}^d$ is a compact set. Then the infimum in the definition of $\overline{\delta}(\eta,\gamma)$ is achieved,*

$$(\eta,\gamma) \mapsto \overline{\delta}(\eta,\gamma) \text{ is narrowly lower semicontinuous} \tag{3.7}$$

and

$$\overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma_h) \le \delta(\boldsymbol{i}, h) \qquad \forall h \in \mathrm{SDiff}(D). \tag{3.8}$$

Furthermore, $\sup_{\eta,\,\gamma\in\Gamma(D)} \overline{\delta}(\eta,\gamma) \le \sqrt{d}$ *when either* $D=[0,1]^d$ *or* $D=\mathbb{T}^d$ *and, more generally,*

$$\sup_{\gamma\in\Gamma(D)} \overline{\delta}_D(\gamma_{\boldsymbol{i}},\gamma) \le \mathrm{Lip}(f) \sup_{\gamma'\in\Gamma(D')} \overline{\delta}_{D'}(\gamma_{\boldsymbol{i}},\gamma')$$

whenever a Lipschitz measure-preserving map $f: D' \to D$ *exists.*

Proof. The inequality $\overline{\delta}(\gamma_{\boldsymbol{i}},\gamma_h) \le \delta(\boldsymbol{i},h)$ simply follows by the fact that any smooth flow g induces a generalized one, with the same action, by the formula $\boldsymbol{\eta}=\Phi_\#\mu_D$, where $\Phi: D\to\tilde{\Omega}(D)$ is the map $x\mapsto(g(\cdot,x),x)$. Assuming that some generalized incompressible flow with a finite action between η and γ exists, the existence of an optimal one follows by the narrow lower semicontinuity of $\boldsymbol{\eta}\mapsto\mathscr{A}_T(\boldsymbol{\eta})$ (because $\omega\mapsto\mathscr{A}_T(\omega)$ is lower semicontinuous in $\Omega(D)$) and by the tightness of minimizing sequences (because $\mathscr{A}_T(\omega)$ is coercive in $\Omega(D)$, by the Ascoli-Arzelà theorem). A similar argument also proves the lower semicontinuity of $(\eta,\gamma)\mapsto\overline{\delta}(\eta,\gamma)$, as the conditions (3.4), (3.5) are stable under narrow convergence (of $\boldsymbol{\eta}$ and η, γ).

When either $D=[0,1]^d$ or $D=\mathbb{T}^d$, it follows by the explicit construction in [12], [24] that $\overline{\delta}(\gamma_{\boldsymbol{i}},\gamma_h)\le\sqrt{d}$ for all $h\in S(D)$; by right invariance (see Proposition 3.4 below) the same estimate holds for $\overline{\delta}(\gamma_f,\gamma_h)$ with $f\in S^i(D)$; by density and lower semicontinuity it extends to $\overline{\delta}(\eta,\gamma)$, with η, $\gamma\in\Gamma(D)$.

Let $f: D'\to D$ be a Lipschitz measure-preserving map and $h\in S(D)$; we claim that it suffices to show the existence of $\gamma'\in\Gamma(D')$ such that $(f\times f)_\#\gamma'=(\boldsymbol{i}\times h)_\#\mu_D$. Indeed, if this is proved, since f naturally induces by left composition a map F from $\tilde{\Omega}(D')$ to $\tilde{\Omega}(D)$ given by $(\omega(t),a)\mapsto(f(\omega(t)),a)$, then to any $\boldsymbol{\eta}\in\Omega(D')$ connecting $\boldsymbol{i}$ to γ' we can associate $F_\#\boldsymbol{\eta}$, which will be a generalized incompressible flow connecting $\boldsymbol{i}$ to h. By the trivial estimate

$$\mathscr{A}_T(F_\#\boldsymbol{\eta}) \le \mathrm{Lip}^2(f)\mathscr{A}_T(\boldsymbol{\eta}),$$

one obtains $\overline{\delta}_D(\gamma_{\boldsymbol{i}},h) \le \mathrm{Lip}(f)\overline{\delta}_{D'}(\gamma_{\boldsymbol{i}},\gamma')$. By density and lower semicontinuity we get the estimate on $\overline{\delta}_D(\gamma_{\boldsymbol{i}},\gamma)$ for all $\gamma\in\Gamma(D)$.

Thus, to conclude the proof, we have to construct γ'. Let us consider the disintegration of $\mu_{D'}$ induced by the map f, that is

$$\mu_{D'} = \int_D \mu_y \, d\mu_D(y) \tag{3.9}$$

where, for μ_D-a.e. y, μ_y is a probability measure in D' concentrated on the compact set $f^{-1}(y)$. We now define γ' as

$$\gamma' := \int_D \mu_y \times \mu_{h(y)} \, d\mu_D(y).$$

Clearly the first marginal of γ' is $\mu_{D'}$; since $h\in S(D)$, changing variables in (3.9) one has $\mu_{D'}=\int_D \mu_{h(y)}\,d\mu_D(y)$, and so also the second marginal of γ' is μ_D. Let us now prove that

$(f\times f)_\#\gamma' = (\boldsymbol{i}\times h)_\#\mu_D$: for any $\phi\in C_b(D\times D)$ we have

$$\begin{aligned}\int_{D\times D}\phi(y,y')\,d(f\times f)_\#\gamma'(y,y') &= \int_{D'\times D'}\phi(f(x),f(x'))\,d\gamma'(x,x')\\ &= \int_D\int_{D'\times D'}\phi(f(x),f(x'))\,d\mu_y(x)\,d\mu_{h(y)}(x')\,d\mu_D(y)\\ &= \int_D \phi(y,h(y))\,d\mu_D(y),\end{aligned}$$

where in the last equality we used that μ_y is concentrated on $f^{-1}(y)$ and $\mu_{h(y)}$ is concentrated on $f^{-1}(h(y))$ for μ_D-a.e. y. □

By (3.1), (3.8) and the narrow lower semicontinuity of $\overline{\delta}(\boldsymbol{i},\cdot)$ we get

$$\overline{\delta}(\gamma_{\boldsymbol{i}},h)\le C\|h-\boldsymbol{i}\|^\alpha_{L^2(D)}\qquad \text{if } h\in S(D),\ D=[0,1]^d,\ d\ge 3. \tag{3.10}$$

We conclude this section by pointing out some additional properties of the metric space $(\Gamma(D),\overline{\delta})$.

Proposition 3.4 $(\Gamma(D),\overline{\delta})$ *is a complete metric space, whose convergence implies narrow convergence. Furthermore, the distance* $\overline{\delta}$ *is right invariant under the action of* $S^{\mathrm{i}}(D)$ *on* $\Gamma(D)$. *Finally,* $\overline{\delta}$*-convergence is strictly stronger than narrow convergence and, as a consequence,* $(\Gamma(D),\overline{\delta})$ *is not compact.*

Proof. We will prove that $\overline{\delta}(\eta,\gamma)\ge W_2(\eta,\gamma)$, where W_2 is the quadratic Wasserstein distance in $\mathscr{P}(D\times D)$ (with the quadratic cost $c((x_1,x_2),(y_1,y_2)) = d_D^2(x_1,y_1)/2 + d_D^2(x_2,y_2)/2$); as this distance metrizes the narrow convergence, this will give the implication between $\overline{\delta}$-convergence and narrow convergence. In order to show the inequality $\overline{\delta}(\eta,\gamma)\ge W_2(\eta,\gamma)$ we consider an optimal flow $\boldsymbol{\eta}_a\otimes\mu_D$ defined in $[0,1]$; then, denoting by $\omega_a\in\Omega(D)$ the constant path identically equal to a, and by $\boldsymbol{\nu}_a\in\mathscr{P}(C([0,1];D\times D))$ the measure $\boldsymbol{\eta}_a\times\delta_{\omega_a}$, the measure $\boldsymbol{\nu} := \int_D \boldsymbol{\nu}_a\,d\mu_D(a)\in\mathscr{P}(C([0,1];D\times D))$ provides a "dynamical transference plan" connecting η to γ (i.e. $(e_0)_\#\boldsymbol{\nu}=\eta$, $(e_1)_\#\boldsymbol{\nu}=\gamma$, see [27, Chapter 7]) whose action is $\overline{\delta}^2(\eta,\gamma)$; since the action of any dynamical transference plan bounds from above $W_2^2(\eta,\gamma)$, the inequality is achieved.

The completeness of $(\Gamma(D),\overline{\delta})$ is a consequence of the inequality $\overline{\delta}\ge W_2$ (so that Cauchy sequences in this space are Cauchy sequences for the Wasserstein distance), the completeness of the Wasserstein spaces of probability measures and the narrow lower semicontinuity of $\overline{\delta}$: we leave the details of the simple proof to the reader.

The right invariance of $\overline{\delta}$ simply follows by the fact that $\eta\circ h = \eta_{h(a)}\otimes\mu_D$, $\gamma\circ h=\gamma_{h(a)}\otimes\mu_D$, so that

$$\overline{\delta}(\eta\circ h,\gamma\circ h)\le\overline{\delta}(\eta,\gamma),$$

because we can apply the same transformation to any admissible flow $\boldsymbol{\eta}_a\otimes\mu_D$ connecting η to γ, producing an admissible flow $\boldsymbol{\eta}_{h(a)}\otimes\mu_D$ between $\eta\circ h$ and $\gamma\circ h$ with the same action. If $h\in S^i(D)$ the inequality can be reversed, using h^{-1}.

Now, let us prove the last part of the statement. We first show that

$$\frac{1}{2}\int_D d_D^2(f,h)\,d\mu_D \le \overline{\delta}^2(\gamma_f,\gamma_h) \qquad \forall f,\, h\in S(D). \tag{3.11}$$

Indeed, considering again an optimal flow $\boldsymbol{\eta}_a\otimes\mu_D$, for μ_D-a.e. $a\in D$ we have

$$\frac{1}{2}d_D^2(f(a),h(a)) = W_2^2(\delta_{f(a)},\delta_{h(a)}) \le T\int_{\Omega(D)}\int_0^T \frac{1}{2}|\dot\omega(t)|^2\,dt\,d\boldsymbol{\eta}_a(\omega),$$

and we need only to integrate this inequality with respect to a. From (3.11) we obtain that $S(D)$ is a closed subset of $\Gamma(D)$, relative to the distance $\overline{\delta}$. In particular, considering for instance a sequence $(g_n)\subset S(D)$ narrowly converging to $\gamma\in\Gamma(D)\setminus S(D)$, whose existence is ensured by (2.6), one proves that the two topologies are not equivalent and the space is not compact. □

Combining right invariance with (3.10), we obtain

$$\overline{\delta}(\gamma_g,\gamma_h) = \overline{\delta}(\gamma_{\boldsymbol{i}},\gamma_{h\circ g^{-1}}) \le C\|g-h\|^\alpha_{L^2(D)} \qquad \forall h\in S(D),\ g\in S^i(D) \tag{3.12}$$

if $D=[0,1]^d$ with $d\ge 3$. By the density of $S^i(D)$ in $S(D)$ in the L^2 norm and the lower semicontinuity of $\overline{\delta}$, this inequality still holds when $g\in S(D)$.

3.3 Brenier's Eulerian-Lagrangian model

In [16], Brenier proposed a second possible relaxation of Arnold's problem, motivated by the fact that this second relaxation allows for a much more precise description of the pressure field, compared to the Lagrangian model (see Section 6).

Still denoting by $\eta=\eta_a\otimes\mu_D\in\Gamma(D)$, $\gamma=\gamma_a\otimes\mu_D\in\Gamma(D)$ the initial and final plan, respectively, the idea is to add to the Eulerian variable x a Lagrangian one a (which, in the case $\eta=\gamma_{\boldsymbol{i}}$, simply labels the position of the particle at time 0) and to consider the family of distributional solutions, indexed by $a\in D$, of the continuity equation

$$\partial_t c_{t,a} + \mathrm{div}(\boldsymbol{v}_{t,a}c_{t,a}) = 0 \qquad \text{in } \mathcal{D}'((0,T)\times D), \quad \text{for } \mu_D\text{-a.e. } a, \tag{3.13}$$

with the initial and final conditions

$$c_{0,a}=\eta_a, \qquad c_{T,a}=\gamma_a, \qquad \text{for } \mu_D\text{-a.e. } a. \tag{3.14}$$

Notice that minimization of the kinetic energy $\int_0^T\int_D |\boldsymbol{v}_{t,a}|^2\,dc_{t,a}\,dt$ among all possible solutions of the continuity equation would give, according to [8], the optimal transport problem between η_a and γ_a (for instance, a path of Dirac masses on a geodesic connecting $g(a)$ to $h(a)$ if $\eta_a=\delta_{g(a)}$, $\gamma_a=\delta_{h(a)}$). Here, instead, by averaging with respect to a we minimize the *mean* kinetic energy

$$\int_D\int_0^T\int_D |\boldsymbol{v}_{t,a}|^2\,dc_{t,a}\,dt\,d\mu_D(a)$$

with the only *global* constraint between the family $\{c_{t,a}\}$ given by the incompressibility of the flow:

$$\int_D c_{t,a}\, d\mu_D(a) = \mu_D \qquad \forall t \in [0,T]. \tag{3.15}$$

It is useful to rewrite this minimization problem in terms of the the global measure c in $[0,T] \times D \times D$ and the measures c_t in $D \times D$

$$c := c_{t,a} \otimes (\mathscr{L}^1 \times \mu_D), \qquad c_t := c_{t,a} \otimes \mu_D$$

(from whom $c_{t,a}$ can obviously be recovered by disintegration), and the velocity field $\boldsymbol{v}(t,x,a) := \boldsymbol{v}_{t,a}(x)$: the action becomes

$$\mathscr{A}_T(c,\boldsymbol{v}) := T \int_0^T \int_{D\times D} \frac{1}{2} |\boldsymbol{v}(t,x,a)|^2 \, dc(t,x,a),$$

while (3.13) is easily seen to be equivalent to

$$\frac{d}{dt} \int_{D\times D} \phi(x,a)\, dc_t(x,a) = \int_{D\times D} \langle \nabla_x \phi(x,a), \boldsymbol{v}(t,x,a)\rangle \, dc_t(x,a) \tag{3.16}$$

for all $\phi \in C_b(D \times D)$ with a bounded gradient with respect to the x variable.

Thus, we can minimize the action on the class of couples measures-velocity fields $(c,\boldsymbol{v})$ that satisfy (3.16) and (3.15), with the endpoint condition (3.14). The existence of a minimum in this class can be proved by standard compactness and lower semicontinuity arguments (see [16] for details). This minimization problem leads to a squared distance between η and γ, that we shall still denote by $\bar{\delta}^2(\eta,\gamma)$. Our notation is justified by the essential equivalence of the two models, proved in the next section.

4 Equivalence of the two relaxed models

In this section we show that the Lagrangian model is equivalent to the Eulerian-Lagrangian one, in the sense that minimal values are the same, and there is a way (not canonical, in one direction) to pass from minimizers of one problem to minimizers of the other one.

Theorem 4.1 *With the notations of Sections 3.2 and 3.3,*

$$\min_{\boldsymbol{\eta}} \mathscr{A}_T(\boldsymbol{\eta}) = \min_{(c,\boldsymbol{v})} \mathscr{A}_T(c,\boldsymbol{v})$$

for any η, $\gamma \in \Gamma(D)$. More precisely, any minimizer $\boldsymbol{\eta}$ of the Lagrangian model connecting η to γ induces in a canonical way a minimizer $(c,\boldsymbol{v})$ of the Eulerian-Lagrangian one, and satisfies for $\mathscr{L}^1$-a.e. $t \in [0,T]$ the condition

$$\dot\omega(t) = \boldsymbol{v}_{t,a}(e_t(\omega)) \qquad \textit{for } \boldsymbol{\eta}\textit{-a.e. } (\omega,a). \tag{4.1}$$

Proof. Up to an isometric embedding, we shall assume that $D \subset \mathbb{R}^m$ isometrically (this is needed to apply Lemma 2.3). If $\boldsymbol{\eta} = \boldsymbol{\eta}_a \otimes \mu \in \mathscr{P}(\tilde{\Omega}(D))$ is a generalized incompressible flow, we denote by $D' \subset D$ a Borel set of full measure such that $\mathscr{A}_T(\boldsymbol{\eta}_a) < \infty$ for all $a \in D'$. For any $a \in D'$ we define

$$c^{\boldsymbol{\eta}}_{t,a} := (e_t)_\# \boldsymbol{\eta}_a, \qquad \boldsymbol{m}^{\boldsymbol{\eta}}_{t,a} = (e_t)_\# \left(\dot{\omega}(t)\boldsymbol{\eta}_a\right).$$

Notice that $\boldsymbol{m}^{\boldsymbol{\eta}}_{t,a}$ is well defined for $\mathscr{L}^1$-a.e. t, and absolutely continuous with respect to $c^{\boldsymbol{\eta}}_{t,a}$, thanks to Lemma 2.3; moreover, denoting by $\boldsymbol{v}^{\boldsymbol{\eta}}_{t,a}$ the density of $\boldsymbol{m}^{\boldsymbol{\eta}}_{t,a}$ with respect to $c^{\boldsymbol{\eta}}_{t,a}$, by the same lemma we have

$$\int_D |\boldsymbol{v}^{\boldsymbol{\eta}}_{t,a}|^2 \, dc^{\boldsymbol{\eta}}_{t,a} \le \int_{\Omega(D)} |\dot{\omega}(t)|^2 \, d\boldsymbol{\eta}_a(\omega), \tag{4.2}$$

with equality only if $\dot{\omega}(t) = \boldsymbol{v}^{\boldsymbol{\eta}}_{t,a}(e_t(\omega))$ for $\boldsymbol{\eta}_a$-a.e. ω. Then, we define the global measure and velocity by

$$c^{\boldsymbol{\eta}} := c^{\boldsymbol{\eta}}_{t,a} \otimes (\mathscr{L}^1 \times \mu_D), \qquad \boldsymbol{v}^{\boldsymbol{\eta}}(t,x,a) = \boldsymbol{v}^{\boldsymbol{\eta}}_t(x,a) := \boldsymbol{v}^{\boldsymbol{\eta}}_{t,a}(x).$$

It is easy to check that $(c^{\boldsymbol{\eta}}, \boldsymbol{v}^{\boldsymbol{\eta}})$ is admissible: indeed, writing $\eta = \eta_a \otimes \mu$, $\gamma = \gamma_a \otimes \mu_D$, the conditions $(e_0)_\# \boldsymbol{\eta}_a = \eta_a$ and $(e_T)_\# \boldsymbol{\eta}_a = \gamma_a$ yield $c^{\boldsymbol{\eta}}_{0,a} = \eta_a$ and $c^{\boldsymbol{\eta}}_{T,a} = \gamma_a$ (for μ_D-a.e. a).

This proves that (3.14) is fulfilled; the incompressibility constraint (3.15) simply comes from (3.4). Finally, we check (3.13) for $a \in D'$; this is equivalent, recalling the definition of $\boldsymbol{v}_{t,a}$, to

$$\frac{d}{dt} \int_D \phi(x) \, dc^{\boldsymbol{\eta}}_{t,a}(x) = \int_D \langle \nabla \phi, \boldsymbol{m}^{\boldsymbol{\eta}}_{t,a} \rangle, \tag{4.3}$$

which in turn corresponds to

$$\frac{d}{dt} \int_{\Omega(D)} \phi(\omega(t)) \, d\boldsymbol{\eta}_a(\omega) = \int_{\Omega(D)} \langle \nabla \phi(\omega(t)), \dot{\omega}(t) \rangle \, d\boldsymbol{\eta}_a(\omega). \tag{4.4}$$

This last identity is a direct consequence of an exchange of differentiation and integral.

By integrating (4.2) in time and with respect to a we obtain that $\mathscr{A}_T(c^{\boldsymbol{\eta}}, \boldsymbol{v}^{\boldsymbol{\eta}}) \le \mathscr{A}_T(\boldsymbol{\eta})$, and equality holds only if (4.1) holds.

So, in order to conclude the proof, it remains to find, given a couple measure-velocity field $(c, \boldsymbol{v})$ with finite action that satisfies (3.13), (3.14) and (3.15), an admissible generalized incompressible flow $\boldsymbol{\eta}$ with $\mathscr{A}_T(\boldsymbol{\eta}) \le \mathscr{A}_T(c, \boldsymbol{v})$. By applying Theorem 2.4 to the family of solutions of the continuity equations (3.13), we obtain probability measures $\boldsymbol{\eta}_a$ with $(e_t)_\# \boldsymbol{\eta}_a = c_{t,a}$ and

$$\int_{\Omega(D)} \int_0^T |\dot{\omega}(t)|^2 \, dt \, d\boldsymbol{\eta}_a(\omega) \le \int_0^T \int_D |\boldsymbol{v}(t,x,a)|^2 \, dc_{t,a}(x) \, dt. \tag{4.5}$$

Then, because of (3.15), it is easy to check that $\boldsymbol{\eta} := \boldsymbol{\eta}_a \otimes \mu_D$ is a generalized incompressible flow, and moreover $\boldsymbol{\eta}$ connects η to γ. By integrating (4.5) with respect to a, we obtain that $\mathscr{A}_T(\boldsymbol{\eta}) \le \mathscr{A}_T(c, \boldsymbol{v})$. $\square$

5 Comparison of metrics and gap phenomena

Throughout this section we shall assume that $D = [0,1]^d$. In [24], Shnirelman proved when $d \geq 3$ the following remarkable approximation theorem for Brenier's generalized (Lagrangian) flows:

Theorem 5.1 *If $d \geq 3$, then each generalized incompressible flow $\boldsymbol{\eta}$ connecting $\boldsymbol{i}$ to $h \in \mathrm{SDiff}(D)$ may be approximated together with the action by a sequence of smooth flows $(g_k(t,\cdot))$ connecting $\boldsymbol{i}$ to h. More precisely:*

(i) the measures $\boldsymbol{\eta}_k := (g_k(\cdot,x))_\# \mu_D$ narrowly converge in $\Omega(D)$ to $\boldsymbol{\eta}$;

(ii) $\mathscr{A}_T(g_k) = \mathscr{A}_T(\boldsymbol{\eta}_k) \to \mathscr{A}_T(\boldsymbol{\eta})$.

This result yields, as a byproduct, the identity

$$\overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma_h) = \delta(\boldsymbol{i}, h) \qquad \text{for all } h \in \mathrm{SDiff}(D),\ d \geq 3. \tag{5.1}$$

More generally the relaxed distance $\overline{\delta}(\eta,\gamma)$ arising from the Lagrangian model can be compared, at least when $\eta = \gamma_{\boldsymbol{i}}$ and the final condition γ is induced by a map $h \in S(D)$, with the relaxation δ_* of the Arnold distance:

$$\delta_*(h) := \inf\left\{\liminf_{n\to\infty} \delta(\boldsymbol{i}, h_n):\ h_n \in \mathrm{SDiff}(D),\ \int_D |h_n - h|^2\, d\mu_D \to 0\right\}. \tag{5.2}$$

By (3.7) and (3.8), we have $\delta_*(h) \geq \overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma_h)$, and a *gap phenomenon* is said to occur if the inequality is strict.

In the case $d = 2$, while examples of $h \in \mathrm{SDiff}(D)$ such that $\delta(\boldsymbol{i}, h) = +\infty$ are known [23], the nature of $\delta_*(h)$ and the possible occurrence of the gap phenomenon are not clear.

In this section we prove the non-occurrence of the gap phenomenon when the final condition belongs to $S(D)$, and even when it is a transport plan, still under the assumption $d \geq 3$. To this aim, we first extend the definition of δ_* by setting

$$\delta_*(\gamma) := \inf\left\{\liminf_{n\to\infty} \delta(\boldsymbol{i}, h_n):\ h_n \in \mathrm{SDiff}(D),\ \gamma_{h_n} \to \gamma \text{ narrowly}\right\}. \tag{5.3}$$

This extends the previous definition (5.2), taking into account that γ_{h_n} narrowly converge to γ_h if and only if $h_n \to h$ in $L^2(\mu_D)$ (for instance, this is a simple consequence of [4, Lemma 2.3]).

Theorem 5.2 *If $d \geq 3$, then $\delta_*(\gamma) = \overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma)$ for all $\gamma \in \Gamma(D)$.*

The proof of the theorem, given at the end of this section, is a direct consequence of Theorem 5.1 and of the following approximation result of generalized incompressible flows by measure-preserving maps (possibly not smooth, or not injective), valid in *any* number of dimensions.

Theorem 5.3 *Let $\gamma \in \Gamma(D)$. Then, for any probability measure $\boldsymbol{\eta}$ on $\Omega(D)$ such that*

$$(e_t)_\# \boldsymbol{\eta} = \mu_D \quad \forall t \in [0,T], \qquad (e_0, e_T)_\# \boldsymbol{\eta} = \gamma,$$

and $\mathscr{A}_T(\boldsymbol{\eta}) < \infty$, there exists a sequence of flows $(g_k(t,\cdot))_{k\in\mathbb{N}} \subset W^{1,2}\big([0,T]; L^2(D)\big)$ such that:

(i) $g_k(t,\cdot) \in S(D)$ *for all* $t \in [0,T]$*, hence* $\boldsymbol{\eta}_k := (\Phi_{g_k})_\# \mu_D$*, with* $\Phi_{g_k}(x) = g_k(\cdot, x)$*, are generalized incompressible flows;*

(ii) $\boldsymbol{\eta}_k$ *narrowly converge in* $\Omega(D)$ *to* $\boldsymbol{\eta}$ *and* $\mathscr{A}_T(g_k) = \mathscr{A}_T(\boldsymbol{\eta}_k) \to \mathscr{A}_T(\boldsymbol{\eta})$*.*

Proof. The first three steps of the proof are more or less the same as in the proof of Shnirelman's approximation theorem (Theorem 5.1 in [24]).

Step 1. Given $\varepsilon > 0$ small, consider the affine transformation of D into the concentric cube D_ε of size $1 - 4\varepsilon$:

$$T_\varepsilon(x) := (2\varepsilon, \dots, 2\varepsilon) + (1 - 4\varepsilon)x.$$

This transformation induces a map $\tilde{T}_\varepsilon$ from $\Omega(D)$ into $C([0,T]; D_\varepsilon)$ (which is indeed a bijection) given by

$$\tilde{T}_\varepsilon(\omega)(t) := T_\varepsilon(\omega(t)) \qquad \forall \omega \in \Omega(D).$$

Then we define $\tilde{\boldsymbol{\eta}}_\varepsilon := (\tilde{T}_\varepsilon)_\# \boldsymbol{\eta}$, and

$$\boldsymbol{\eta}_\varepsilon := (1 - 4\varepsilon)^d \tilde{\boldsymbol{\eta}}_\varepsilon + \boldsymbol{\eta}_{0,\varepsilon},$$

where $\boldsymbol{\eta}_{0,\varepsilon}$ is the "steady" flow in $D \setminus D_\varepsilon$: it consists of all the curves in $D \setminus D_\varepsilon$ that do not move for $0 \le t \le T$. It is then not difficult to prove that $\boldsymbol{\eta}_\varepsilon \to \boldsymbol{\eta}$ narrowly and $\mathscr{A}_T(\boldsymbol{\eta}_\varepsilon) \to \mathscr{A}_T(\boldsymbol{\eta})$, as $\varepsilon \to 0$.

Therefore, by a diagonal argument, it suffices to prove our theorem for a measure $\boldsymbol{\eta}$ which is steady near ∂D. More precisely we can assume that, if $\omega(0)$ is in the 2ε-neighborhood of ∂D, then $\omega(t) \equiv \omega(0)$ for $\boldsymbol{\eta}$-a.e. ω. Moreover, arguing as in Step 1 of the proof of the above mentioned approximation theorem in [24], we can assume that the flow does not move for $0 \le t \le \varepsilon$, that is, for $\boldsymbol{\eta}$-a.e. ω, $\omega(t) \equiv \omega(0)$ for $0 \le t \le \varepsilon$.

Step 2. Let us now consider a family of independent random variables $\omega_1, \omega_2, \dots$ defined in a common probability space $(Z, \mathcal{Z}, P)$, with values in $C([0,T], D)$ and having the same law $\boldsymbol{\eta}$. Recall that $\boldsymbol{\eta}$ is steady near ∂D and for $0 \le t \le \varepsilon$, so we can see ω_i as random variables with values in the subset of $\Omega(D)$ given by the curves which do not move for $0 \le t \le \varepsilon$ and in the 2ε-neighbourhood of the ∂D. By the law of large numbers, the random probability measures in $\Omega(D)$

$$\boldsymbol{\nu}_N(z) := \frac{1}{N} \sum_{i=1}^{N} \delta_{\omega_i(z)}, \qquad z \in Z,$$

narrowly converge to $\boldsymbol{\eta}$ with probability 1. Moreover, always by the law of large numbers, also

$$\mathscr{A}_T(\boldsymbol{\nu}_N(z)) \to \mathscr{A}_T(\boldsymbol{\eta})$$

with probability 1. Thus, choosing properly z, we have approximated $\boldsymbol{\eta}$ with measures $\boldsymbol{\nu}_N$ concentrated on a finite number of trajectories $\omega_i(z)(\cdot)$ which are steady in $[0, \varepsilon]$ and close to ∂D. From now on (as typical in Probability theory) the parameter z will be tacitly understood.

Step 3. Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ be a smooth radial convolution kernel with $\varphi(x) = 0$ for $|x| \ge 1$ and $\varphi(x) > 0$ for $|x| < 1$. Given a finite number of trajectories $\omega_1, \dots, \omega_N$ as described is step 2, we define

$$a_i(x) := \frac{1}{\varepsilon^d} \varphi\left(\frac{x - \omega_i(0)}{\varepsilon}\right) \qquad \text{if } \operatorname{dist}(\omega_i(0), \partial D) \ge \varepsilon,$$

$$a_i(x) := \frac{1}{\varepsilon^d}\sum_{\gamma\in\Gamma}\varphi\left(\frac{x-\gamma(\omega_i(0))}{\varepsilon}\right) \qquad \text{if } \operatorname{dist}(\omega_i(0),\partial D)\le\varepsilon,$$

where Γ is the discrete group of motions in $\mathbb{R}^n$ generated by the reflections in the faces of D. It is easy to check that $\int a_i = 1$ and that $\operatorname{supp}(a_i)$ is the intersection of D with the closed ball $\overline{B}_\varepsilon(\omega_i(0))$. Define

$$g_{i,t}(x) := \omega_i(t) + (x-\omega_i(0)) \qquad \forall i=1,\dots,n.$$

Let $\mathscr{M}_N := (a_1,\dots,a_N,g_{1,t}(x),\dots,g_{N,t}(x))$ and let us consider the generalized flow $\boldsymbol{\eta}_N$ associated to $\mathscr{M}_N$, given by

$$\int_{\Omega(D)} f(\omega)\,d\boldsymbol{\eta}_N := \frac{1}{N}\sum_{i=1}^N\int_D a_i(x)f(t\mapsto g_{i,t}(x))\,dx \tag{5.4}$$

(that is, $\boldsymbol{\eta}_N$ is the measure in the space of paths given by $\frac{1}{N}\sum_i\int_D a_i(x)\delta_{g_{i,\cdot}(x)}dx$). The measure $\boldsymbol{\eta}_N$ is well defined for the following reason: if $\operatorname{dist}(\omega_i(0),\partial D)\le\varepsilon$ we have $g_{i,t}(x)=x$, and if $\operatorname{dist}(\omega_i(0),\partial D)>\varepsilon$ and $a_i(x)>0$ we still have that the curve $t\mapsto g_{i,t}(x)$ is contained in D because $a_i(x)>0$ implies $|x-\omega_i(0)|\le\varepsilon$ and, by construction, $\operatorname{dist}(\omega_i(t),\partial D)\ge\varepsilon$ for all times. Since the density $\rho^{\boldsymbol{\eta}_N}$ induced by $\boldsymbol{\eta}_N$ is given by

$$\rho^N(t,x) := \frac{1}{N}\sum_{i=1}^N a_i(x+\omega_i(0)-\omega_i(t)),$$

the flow $\boldsymbol{\eta}_N$ is not measure preserving. However we are more or less in the same situation as in Step 3 in the proof of the approximation theorem in [24] (the only difference being that we do not impose any final data). Thus, by [24, Lemma 1.2], with probability 1

$$\begin{aligned}
&\sup_{x,t}|\rho^N(t,x)-1|\to 0,\\
&\sup_{x,t}|\partial_x^\alpha\rho^N(t,x)|\to 0 \quad \forall\alpha,\\
&\int_D\int_0^T|\partial_t\rho^N(t,x)|^2\,dt\,dx\to 0
\end{aligned} \tag{5.5}$$

as $N\to\infty$. By the first two equations in (5.5), we can left compose $g_{i,t}$ with a smooth correcting flow $\zeta_t^N(x)$ as in Step 3 in the proof of the approximation theorem in [24], in such a way that the flow $\tilde{\boldsymbol{\eta}}_N$ associated to $\tilde{\mathscr{M}}_N := (a_1,\dots,a_N,\zeta_t^N\circ g_{1,t}(x),\dots,\zeta_t^N\circ g_{N,t}(x))$ via the formula analogous to (5.4) is incompressible. Moreover, thanks to the third equation in (5.5) and the convergence of $\mathscr{A}_T(\boldsymbol{\nu}_N)$ to $\mathscr{A}_T(\boldsymbol{\eta})$, one can prove that $\mathscr{A}_T(\tilde{\boldsymbol{\eta}}_N)\to\mathscr{A}_T(\boldsymbol{\eta})$ with probability 1.

We observe that, since $\boldsymbol{\eta}$ is steady for $0\le t\le\varepsilon$, the same holds by construction for $\tilde{\boldsymbol{\eta}}_N$. Without loss of generality, we can therefore assume that ζ_t^N does not depend on t for $t\in[0,\varepsilon]$.

Step 4. In order to conclude, we see that the only problem now is that the flow $\tilde{\boldsymbol{\eta}}_N$ associated to $\tilde{\mathscr{M}}_N$ is still non-deterministic, since if $x\in\operatorname{supp}(a_i)\cap\operatorname{supp}(a_j)$ for $i\ne j$, then more that one curve starts from x. Let us partition D in the following way:

$$D = D_1\cup D_2\cup\ldots\cup D_L\cup E,$$

where E is $\mathscr{L}^d$-negligible, any set D_j is open, and all $x \in D_j$ belong to the interior of the supports of exactly $M = M(j) \leq N$ sets a_i, indexed by $1 \leq i_1 < \cdots < i_M \leq N$ (therefore $L \leq 2^N$). This decomposition is possible, as E is contained in the union of the boundaries of $\operatorname{supp} a_i$, which is $\mathscr{L}^d$-negligible.

Fix one of the sets D_j and assume just for notational simplicity that $i_k = k$ for $1 \leq k \leq M$. We are going to modify the flow $\tilde{\boldsymbol{\eta}}_N$ in D_j, increasing a little bit its action (say, by an amount $\alpha > 0$), in such a way that for each point in D_j only one curve starts from it. Given $x \in D_j$, we know that M curves start from it, weighted with mass $a_k(x) > 0$, and $\sum_{k=1}^M a_k(x) = 1$. These curves coincide for $0 \leq t \leq \varepsilon$ (since nothing moves), and then separate. We want to partition D_j in M sets E_k, with

$$\mathscr{L}^d(E_k) = \int_{D_j} a_k(x)\,dx, \qquad 1 \leq k \leq M$$

in such a way that, for any $x \in E_k$, only one curve ω_x^k starts from it at time 0, $\omega_x^k(t) \in D_j$ for $0 \leq t \leq \varepsilon$, and the map $E_k \ni x \mapsto \omega_x^k(\varepsilon) \in D_j$ pushes forward $\mathscr{L}^d \llcorner E_k$ into $a_k \mathscr{L}^d \llcorner D_j$. Moreover, we want the incompressibility condition to be preserved for all $t \in [0, \varepsilon]$. If this is possible, the proof will be concluded by gluing ω_x^k with the only curve starting from $\omega_x^k(\varepsilon)$ with weight $a_k(\omega^k(\varepsilon))$.

The above construction can be achieved in the following way. First we write the interior of D_j, up to null measure sets, as a countable union of disjoints open cubes (C_i) with size δ_i satisfying

$$\frac{M^2}{\varepsilon} \sum_i \frac{\delta_i^2}{\bar{b}_i^2} \mathscr{L}^d(C_i) \leq \alpha, \tag{5.6}$$

with $\bar{b}_i := \min_{1 \leq k \leq M} \min_{C_i} a_k$. This is done just considering the union of the grids in $\mathbb{R}^d$ given by $\mathbb{Z}^d/2^n$ for $n \in \mathbb{N}$, and taking initially our cubes in this family; if (5.6) does not hold, we keep splitting the cubes until it is satisfied ($\bar{b}_i$ can only increase under this additional splitting, therefore a factor 4 is gained in each splitting). Once this partition is given, the idea is to move the mass within each C_i for $0 \leq t \leq \varepsilon$. At least heuristically, one can imagine that in C_i the functions a_k are almost constant and that the velocity of a generic path in C_i is at most of order δ_i/ε. Thus, the total energy of the new incompressible fluid in the interval $[0, \varepsilon]$ will be of order

$$\sum_i \int_{C_i} \int_0^\varepsilon |\dot{\omega}_x(t)|^2\,dt\,dx \leq \frac{C}{\varepsilon} \sum_i \delta_i^2 \mathscr{L}^d(C_i)$$

and the conclusion will follow by our choice of δ_i.

So, in order to make this argument rigorous, let us fix i and let us see how to construct our modified flow in the cube C_i for $t \in [0, \varepsilon]$. Slicing C_i with respect to the first $(d-1)$-variables, we see that the transport problem can be solved in each slice. Specifically, if C_i is of the form $x^i + (0, \delta_i)^d$, and we define

$$m^k := \int_{C_i} a_k(x)\,dx, \qquad k = 1, \ldots, M,$$

whose sum is δ_i^d, then the points which belong to $C_i^k := x^i + (0,\delta_i)^{d-1} \times J_k$ have to move along curves in order to push forward $\mathscr{L}^d \llcorner C_i^k$ into $a_k \mathscr{L}^d \llcorner C_i$, where J_k are M consecutive open intervals in $(0,\delta_i)$ with length $\delta_i^{1-d} m^k$. Moreover, this has to be done preserving the incompressibility condition.

If we write $x = (x', x_d) \in \mathbb{R}^d$ with $x' = (x_1, \ldots, x_{d-1})$, we can transport the M uniform densities

$$\mathscr{H}^1 \llcorner \big(x^i + \{x'\} \times J_k\big) \quad \text{with } x' \in [0,\delta_i]^{(d-1)},$$

into the M densities

$$a_k(x', \cdot) \mathscr{H}^1 \llcorner \big(x^i + \{x'\} \times [0,\delta_i]\big)$$

moving the curves only in the d-th direction, i.e. keeping x' fixed. Thanks to Lemma 5.4 below and a scaling argument, we can do this construction paying at most $M^2 \bar{b}_i^{-2} \delta_i^3 / \varepsilon$ in each slice of C_i, and therefore with a total cost less than

$$\frac{M^2}{\varepsilon} \sum_i \frac{\delta_i^{d+2}}{\bar{b}_i^2} \leq \alpha.$$

This concludes our construction. □

Lemma 5.4 *Let $M \geq 1$ be an integer and let $b_1, \ldots, b_M : [0,1] \to (0,1]$ be continuous with $\sum_1^M b_k = 1$. Setting $l_k = \int_0^1 b_k \, dt \in (0,1]$, and denoting by $J_1, \ldots, J_M$ consecutive intervals of $(0,1)$ with length l_k, there exists a family of uniformly Lipschitz maps $h(\cdot, x)$, with $h(t, \cdot) \in S([0,1])$, such that*

$$h(1,\cdot)_\#(\chi_{J_k} \mathscr{L}^1) = b_k \mathscr{L}^1, \qquad k = 1, \ldots, M$$

and

$$\mathscr{A}_1(h) \leq \frac{M^2}{\bar{b}^2}, \qquad \text{with } \bar{b} := \min_{1 \leq k \leq M} \min_{[0,1]} b_k > 0. \tag{5.7}$$

Proof. We start with a preliminary remark: let $J \subset (0,1)$ be an interval with length l and assume that $t \mapsto \rho_t$ is a nonnegative Lipschitz map between $[0,1]$ and $L^1(0,1)$, with $\rho_t \leq 1$ and $\int_0^1 \rho_t \, dx = l$ for all $t \in [0,1]$, and let $f(t,\cdot)$ be the unique (on J, up to countable sets) nondecreasing map pushing $\chi_J \mathscr{L}^1$ to ρ_t. Assume also that $\operatorname{supp} \rho_t$ is an interval and $\rho_t \geq r$ $\mathscr{L}^1$-a.e. on $\operatorname{supp} \rho_t$, with $r > 0$. Under this extra assumption, $f(t,x)$ is uniquely determined for all $x \in J$, and implicitly characterized by the conditions

$$\int_0^{f(t,x)} \rho_t(y) \, dy = \mathscr{L}^1((0,x) \cap J), \qquad f(t,x) \in \operatorname{supp} \rho_t.$$

This implies, in particular, that $f(\cdot, x)$ is continuous *for all* $x \in J$. We are going to prove that this map is even Lipschitz continuous in $[0,1]$ and

$$\left|\frac{d}{dt} f(t,x)\right| \leq \frac{\mathrm{Lip}(\rho_\cdot)}{r} \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in [0,1] \tag{5.8}$$

for all $x \in J$. To prove this fact, we first notice that the endpoints of the interval $\operatorname{supp}\rho_t$ (whose length is at least l) move at most with velocity $\mathrm{Lip}(\rho_\cdot)/r$; then, we fix $x \in J = [a,b]$ and consider separately the cases

$$x \in \partial J = \{a,b\}, \qquad x \in \mathrm{Int}(J) = (a,b).$$

In the first case, since for any $t \in [0,1]$

$$\int_0^{f(t,a)} \rho_t(y)\,dy = 0, \qquad \int_0^{f(t,b)} \rho_t(y)\,dy = \mathscr{L}^1(J),$$

and by assumption $f(t,x) \in \operatorname{supp}\rho_t$ for any $x \in J$, we get $\operatorname{supp}\rho_t = [f(t,a), f(t,b)]$ for all $t \in [0,1]$. This, together with the fact that the endpoints of the interval $\operatorname{supp}\rho_t$ move at most with velocity $\mathrm{Lip}(\rho_\cdot)/r$, implies (5.8) if $x \in \partial J$. In the second case we have

$$\int_0^{f(t,x)} \rho_t(y)\,dy \in (0, \mathscr{L}^1(J)),$$

therefore $f(t,x) \in \mathrm{Int}(\operatorname{supp}\rho_t)$ for all $t \in [0,1]$. It suffices now to find a Lipschitz estimate of $|f(s,x) - f(t,x)|$ when s, t are sufficiently close. Assume that $f(s,x) \le f(t,x)$: adding and subtracting $\int_0^{f(s,x)} \rho_t(y)\,dy$ in the identity

$$\int_0^{f(t,x)} \rho_t(y)\,dy = \int_0^{f(s,x)} \rho_s(y)\,dy$$

we obtain

$$\int_{f(s,x)}^{f(t,x)} \rho_t(y)\,dy = \int_0^{f(s,x)} \rho_s(y) - \rho_t(y)\,dy.$$

Now, as $f(s,x)$ belongs to $\operatorname{supp}\rho_t$ for $|s-t|$ sufficiently small, we get

$$r|f(s,x) - f(t,x)| \le \mathrm{Lip}(\rho_\cdot)|t-s|.$$

This proves the Lipschitz continuity of $f(\cdot,x)$ and (5.8).

Given this observation, to prove the lemma it suffices to find maps $t \mapsto \rho_t^k$ connecting $\chi_{J_k}\mathscr{L}^1$ to $b_k\mathscr{L}^1$ satisfying:

(i) $\operatorname{supp}\rho_t^k$ is an interval, and $\rho_t^k \ge \min_{[0,1]} b_k \ge \bar{b}$ $\mathscr{L}^1$-a.e. on its support;

(ii) $\mathrm{Lip}(\rho_\cdot^k) \le \frac{M-1}{2}$ on $[0,\frac{1}{2}]$, and $\mathrm{Lip}(\rho_\cdot^k) \le 2$ on $[\frac{1}{2},1]$;

(iii) $\sum_{k=1}^{M} \rho_t^k = 1$ for all $t \in [0,1]$.

Indeed, this would produce maps with time derivative bounded by $(M-1)/(2\bar{b})$ on $[0,\frac{1}{2}]$ and bounded by $2/\bar{b}$ on $[\frac{1}{2},1]$, and this easily gives (5.7).

The construction can be achieved in two steps. First, we connect $\chi_{J_k}\mathscr{L}^1$ to $l_k\mathscr{L}^1$ in the time interval $[0,\frac{1}{2}]$; then, we connect $l_k\mathscr{L}^1$ to $b_k\mathscr{L}^1$ in $[\frac{1}{2},1]$ by a linear interpolation. The

Lipschitz constants of the second step are easily seen to be less than 2, so let us focus on the first interpolation.

Let us first consider the case of two densities $\rho^1 = \chi_{J_1}$ and $\rho^2 = \chi_{J_2}$, with $J_1 = (0, l_1)$ and $J_2 = (l_1, l)$. In the time interval $[0, \tau]$, we define the expanding intervals

$$J_{1,t} = (0, l_1 + \frac{t}{\tau} l_2), \quad J_{2,t} = (l_1 - \frac{t}{\tau} l_1, 1),$$

so that $J_{k,\tau} = (0, l)$ for $k = 1, 2$, and then define

$$\rho_t^1 := \begin{cases} 1 & \text{on } (0, l_1 - \frac{t}{\tau} l_1), \\ l_1/l & \text{on } (l_1 - \frac{t}{\tau} l_1, l_1 + \frac{t}{\tau} l_2), \\ 0 & \text{otherwise.} \end{cases} \qquad \rho_t^2 := \begin{cases} 1 & \text{on } (l_1 + \frac{t}{\tau} l_2, l), \\ l_2/l & \text{on } (l_1 - \frac{t}{\tau} l_1, l_1 + \frac{t}{\tau} l_2), \\ 0 & \text{otherwise.} \end{cases}$$

By construction $\rho_t^k \geq l_k$ on $J_{k,t}$ for $k = 1, 2$, $\rho_t^1 + \rho_t^2 = 1$, and it is easy to see that

$$\mathrm{Lip}(\rho_\cdot^k) \leq \frac{l_1 l_2}{\tau l} \leq \frac{l}{4\tau}. \tag{5.9}$$

We can now define the desired interpolation on $[0, \frac{1}{2}]$ for general $M \geq 2$. Let us define

$$t_i := \frac{i}{2(M-1)} \qquad \text{for } i = 1, \ldots, M-1,$$

so that $t_{M-1} = \frac{1}{2}$. We will achieve our construction of ρ_t^k on $[0, \frac{1}{2}]$ in $M-1$ steps, where at each step we will progressively define ρ_t^k on the time interval $[t_{i-1}, t_i]$.

First, in the time interval $[0, t_1]$, we leave fixed $\rho_0^k := \chi_{J_k} \mathscr{L}^1$ for $k \geq 3$ (if such k exist), while we apply the above construction in $J_1 \cup J_2$ to ρ^1 and ρ^2. In this way, on $[0, t_1]$, $\rho_0^1 := \chi_{J_1} \mathscr{L}^1$ is connected to $\rho_{t_1}^1 := \frac{l_1}{l_1 + l_2} \chi_{J_1 \cup J_2} \mathscr{L}^1$, and $\rho_0^2 := \chi_{J_2} \mathscr{L}^1$ is connected to $\rho_{t_1}^2 := \frac{l_2}{l_1 + l_2} \chi_{J_1 \cup J_2} \mathscr{L}^1$.

Now, as a second step, we want to connect $\rho_{t_1}^k$ to $\frac{l_k}{l_1 + l_2 + l_3} \chi_{J_1 \cup J_2 \cup J_3} \mathscr{L}^1$ for $k = 1, 2, 3$, leaving the other densities fixed. To this aim, we define $\rho_{t_1}^{12} := \rho_{t_1}^1 + \rho_{t_1}^2 = \chi_{J_1 \cup J_2} \mathscr{L}^1$. In the time interval $[t_1, t_2]$, we leave fixed $\rho_0^k := \chi_{J_k} \mathscr{L}^1$ for $k \geq 4$ (if such k exist), and we apply again the above construction in $J_1 \cup J_2 \cup J_3$ to $\rho_{t_1}^{12}$ and $\rho_{t_1}^3 = \chi_{J_3} \mathscr{L}^1$. In this way, on $[t_1, t_2]$, $\rho_{t_1}^{12}$ is connected to $\rho_{t_2}^{12} := \frac{l_1 + l_2}{l_1 + l_2 + l_3} \chi_{J_1 \cup J_2 \cup J_3} \mathscr{L}^1$, and $\rho_{t_1}^3$ is connected to $\rho_{t_2}^3 := \frac{l_3}{l_1 + l_2 + l_3} \chi_{J_1 \cup J_2 \cup J_3} \mathscr{L}^1$. Finally, it suffices to define $\rho_t^1 := \frac{l_1}{l_1 + l_2} \rho_t^{12}$ and $\rho_t^2 := \frac{l_2}{l_1 + l_2} \rho_t^{12}$.

In the third step we leave fixed the densities $\rho_{t_2}^k$ for $k \geq 5$, and we do the same construction as before adding the first three densities (that is, in this case one defines $\rho_{t_2}^{123} := \rho_{t_2}^1 + \rho_{t_2}^2 + \rho_{t_2}^3 = \chi_{J_1 \cup J_2 \cup J_3} \mathscr{L}^1$). In this way, we connect $\rho_{t_2}^{123}$ to $\rho_{t_3}^{123} := \frac{l_1 + l_2 + l_3}{l_1 + l_2 + l_3 + l_4} \chi_{J_1 \cup J_2 \cup J_3 \cup J_4} \mathscr{L}^1$ and $\rho_{t_2}^4$ to $\rho_{t_3}^4 := \frac{l_4}{l_1 + l_2 + l_3 + l_4} \chi_{J_1 \cup J_2 \cup J_3 \cup J_4} \mathscr{L}^1$, and then we define $\rho_t^k := \frac{l_k}{l_1 + l_2 + l_3} \rho_t^{123}$ for $k = 1, 2, 3$.

Iterating this construction on $[t_i, t_{i+1}]$ for $i \geq 4$, one obtains the desired maps $t \mapsto \rho_t^k$. Indeed, by construction $\rho_t^k \geq l_k$ on $J_{k,t}$, and $\sum_{k=1}^M \rho_t^k = 1$. Moreover, by (5.9), it is simple to see that in each time interval $[t_i, t_{i+1}]$ one has the bound

$$\mathrm{Lip}(\rho_\cdot^k) \leq \frac{M-1}{2}.$$

So the energy can be easily bounded by $1/\bar{b}^2 \big(\frac{(M-1)^2}{16} + 1 \big) \leq M^2/\bar{b}^2$. □

Proof. (of Theorem 5.2) By applying Theorem 5.3 to the optimal $\boldsymbol{\eta}$ connecting $\boldsymbol{i}$ to γ, we can find maps $g_k \in S(D)$ such that $\gamma_{g_k} \to \gamma$ narrowly and

$$\limsup_{k\to\infty} \overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma_{g_k}) \le \overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma).$$

Now, if $d \ge 3$ we can use (3.12), the triangle inequality, and the density of $\mathrm{SDiff}(D)$ in $S(D)$ in the L^2 norm, to find maps $h_k \in \mathrm{SDiff}(D)$ such that

$$\limsup_{k\to\infty} \overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma_{h_k}) \le \overline{\delta}(\gamma_{\boldsymbol{i}}, \gamma)$$

and $\gamma_{h_k} \to \gamma$ narrowly. This gives the thesis. □

6 Necessary and sufficient optimality conditions

In this section we study necessary and sufficient optimality conditions for the generalized geodesics; we shall work mainly with the Lagrangian model, but we will use the equivalent Eulerian-Lagrangian model to transfer regularity informations for the pressure field to the Lagrangian model. Without any loss of generality, we assume throughout this section that $T = 1$.

The pressure field p can be identified, at least as a distribution (precisely, an element of the dual of $C^1([0,1]\times D)$), by the so-called dual least action principle introduced in [14]. In order to describe it, let us build a natural class of first variations in the Lagrangian model: given a smooth vector field $\boldsymbol{w}(t,x)$, vanishing for t sufficiently close to 0 and 1, we may define the maps $S^\varepsilon : \tilde{\Omega}(D) \to \tilde{\Omega}(D)$ by

$$S^\varepsilon(\omega, a)(t) := (e^{\varepsilon \boldsymbol{w}_t}\omega(t), a)\,, \tag{6.1}$$

where $e^{\varepsilon \boldsymbol{w}_t}x$ is the flow, in the (ε, x) variables, generated by the autonomous field $\boldsymbol{w}_t(x) = \boldsymbol{w}(t,x)$ (i.e. $e^{0\boldsymbol{w}_t} = \boldsymbol{i}$ and $\frac{d}{d\varepsilon}e^{\varepsilon \boldsymbol{w}_t}x = \boldsymbol{w}(t, e^{\varepsilon \boldsymbol{w}_t}x)$), and the perturbed generalized flows $\boldsymbol{\eta}_\varepsilon := (S^\varepsilon)_\#\boldsymbol{\eta}$. Notice that $\boldsymbol{\eta}_\varepsilon$ is incompressible if $\mathrm{div}\,\boldsymbol{w}_t = 0$, and more generally the density $\rho^{\boldsymbol{\eta}_\varepsilon}$ satisfies for all times $t \in (0,1)$ the continuity equation

$$\frac{d}{d\varepsilon}\rho^{\boldsymbol{\eta}_\varepsilon}(t,x) + \mathrm{div}(\boldsymbol{w}_t(x)\rho^{\boldsymbol{\eta}_\varepsilon}(t,x)) = 0. \tag{6.2}$$

This motivates the following definition.

Definition 6.1 (Almost incompressible flows) *We say that a probability measure $\boldsymbol{\nu}$ on $\Omega(D)$ is a almost incompressible generalized flow if $\rho^{\boldsymbol{\nu}} \in C^1([0,1]\times D)$ and*

$$\|\rho^{\boldsymbol{\nu}} - 1\|_{C^1([0,1]\times D)} \le \frac{1}{2}.$$

Now we provide a slightly simpler proof of the characterization given in [14] of the pressure field (the original proof therein involved a time discretization argument).

Theorem 6.2 *For all η, $\gamma \in \Gamma(D)$ there exists $p \in \left[C^1([0,1] \times D)\right]^*$ such that*

$$\langle p, \rho^{\boldsymbol{\nu}} - 1 \rangle_{(C^1)^*, C^1} \leq \mathscr{A}_1(\boldsymbol{\nu}) - \overline{\delta}^2(\eta, \gamma) \tag{6.3}$$

for all almost incompressible flows $\boldsymbol{\nu}$ satisfying (3.5).

Proof. Let us define the closed convex set $C := \{\rho \in C^1([0,1] \times D) : \ \|\rho - 1\|_{C^1} \leq \frac{1}{2}\}$, and the function $\phi : C^1([0,1] \times D) \to \mathbb{R}^+ \cup \{+\infty\}$ given by

$$\phi(\rho) := \begin{cases} \inf\{\mathscr{A}_1(\boldsymbol{\nu}) : \ \rho^{\boldsymbol{\nu}} = \rho \text{ and (3.5) holds}\} & \text{if } \rho \in C; \\ +\infty & \text{otherwise.} \end{cases}$$

We observe that $\phi(1) = \overline{\delta}^2(\eta, \gamma)$. Moreover, it is a simple exercise to prove that ϕ is convex and lower semicontinuous in $C^1([0,1] \times D)$. Let us now prove that ϕ has bounded (descending) slope at 1, i.e.

$$\limsup_{\rho \to 1} \frac{[\phi(1) - \phi(\rho)]^+}{\|1 - \rho\|_{C^1}} < +\infty,$$

By [14, Proposition 2.1] we know that there exist $0 < \varepsilon < \frac{1}{2}$ and $c > 0$ such that, for any $\rho \in C$ with $\|\rho - 1\|_{C^1} \leq \varepsilon$, there is a Lipschitz family of diffeomorphisms $g_\rho(t, \cdot) : D \to D$ such that

$$g_\rho(t, \cdot)_\# \mu_D = \rho(t, \cdot)\mu_D,$$

$g_\rho(t, \cdot) = \boldsymbol{i}$ for $t = 0$, 1, and the Lipschitz constant of $(t, x) \mapsto g_\rho(t, x) - x$ is bounded by c. Thus, adapting the construction in [14, Proposition 2.1] (made for probability measures in $\Omega(D)$, and not in $\tilde{\Omega}(D)$), for any incompressible flow $\boldsymbol{\eta}$ connecting η to γ, and any $\rho \in C$, we can define an almost incompressible flow $\boldsymbol{\nu}$ still connecting η to γ such that $\rho^{\boldsymbol{\nu}} = \rho$, and

$$\mathscr{A}_1(\boldsymbol{\nu}) \leq \mathscr{A}_1(\boldsymbol{\eta}) + c' \|\rho - 1\|_{C^1}(1 + \mathscr{A}_1(\boldsymbol{\eta})),$$

where c' depends only on c (for instance, we define $\boldsymbol{\nu} := G_\# \boldsymbol{\eta}$, where $G : \tilde{\Omega}(D) \to \tilde{\Omega}(D)$ is the map induced by g_ρ via the formula $(\omega(t), a) \mapsto (g_\rho(t, \omega(t)), a)$). In particular, considering an optimal $\boldsymbol{\eta}$, we get

$$\phi(\rho) \leq \phi(1) + c\|\rho - 1\|_{C^1}(1 + \overline{\delta}^2(\eta, \gamma)) \tag{6.4}$$

for any $\rho \in C$ with $\|\rho - 1\|_{C^1} \leq \varepsilon$. This fact implies that ϕ is bounded on a neighbourhood of 1 in C. Now, it is a standard fact of convex analysis that a convex function bounded on a convex set is locally Lipschitz on that set. This provides the bounded slope property. By a simple application of the Hahn-Banach theorem (see for instance Proposition 1.4.4 in [3]), it follows that the subdifferential of ϕ at 1 is not empty, that is, there exists p in the dual of C^1 such that

$$\langle p, \rho - 1 \rangle_{(C^1)^*, C^1} \leq \phi(\rho) - \phi(1).$$

This is indeed equivalent to (6.3). □

This result tells us that, if $\boldsymbol{\eta}$ is an optimal incompressible generalized flow connecting η to γ (i.e. $\mathscr{A}_1(\boldsymbol{\eta}) = \overline{\delta}^2(\eta,\gamma)$), and if we consider the augmented action

$$\mathscr{A}_1^p(\boldsymbol{\nu}) := \int_{\tilde\Omega(D)} \int_0^1 \frac{1}{2}|\dot\omega(t)|^2 \, dt \, d\boldsymbol{\nu}(\omega,a) - \langle p, \rho_{\boldsymbol{\nu}} - 1\rangle, \tag{6.5}$$

then $\boldsymbol{\eta}$ minimizes the new action among all almost incompressible flows $\boldsymbol{\nu}$ between η and γ.

Then, using the identities

$$\frac{d}{d\varepsilon}\frac{d}{dt}S^\varepsilon(\omega)(t)\Big|_{\varepsilon=0} = \frac{d}{dt}\boldsymbol{w}(t,\omega(t)) = \partial_t \boldsymbol{w}(t,\omega(t)) + \nabla_x \boldsymbol{w}(t,\omega(t))\cdot\dot\omega(t)$$

and the convergence in the sense of distributions (ensured by (6.2)) of $(\rho^{\boldsymbol{\eta}_\varepsilon} - 1)/\varepsilon$ to $-\mathrm{div}\,\boldsymbol{w}$ as $\varepsilon \downarrow 0$, we obtain

$$0 = \frac{d}{d\varepsilon}\mathscr{A}_1^p(\boldsymbol{\eta}_\varepsilon)\Big|_{\varepsilon=0} = \int_{\tilde\Omega(D)}\int_0^1 \dot\omega(t)\cdot\frac{d}{dt}\boldsymbol{w}(t,\omega(t))\, dt\, d\boldsymbol{\eta}(\omega,a) + \langle p, \mathrm{div}\,\boldsymbol{w}\rangle. \tag{6.6}$$

As noticed in [14], this equation identifies uniquely the pressure field p (as a distribution) up to trivial modifications, i.e. additive perturbations depending on time only.

In the Eulerian-Lagrangian model, instead, the pressure field is defined (see (2.20) in [16]) and uniquely determined, still up to trivial modifications, by

$$\nabla p(t,x) = -\partial_t\left(\int_D \boldsymbol{v}(t,x,a)\, dc_{t,x}(a)\right) - \mathrm{div}\left(\int_D \boldsymbol{v}(t,x,a)\otimes\boldsymbol{v}(t,x,a)\, dc_{t,x}(a)\right), \tag{6.7}$$

all derivatives being understood in the sense of distributions in $(0,1)\times D$ (here $(c,\boldsymbol{v})$ is any optimal pair for the Eulerian-Lagrangian model). We used the same letter p to denote the pressure field in the two models: indeed, we have seen in the proof of Theorem 4.1 that, writing $\boldsymbol{\eta} = \boldsymbol{\eta}_a \otimes \mu_D$, the correspondence

$$\boldsymbol{\eta} \mapsto (c^{\boldsymbol{\eta}}_{t,a}, \boldsymbol{v}^{\boldsymbol{\eta}}_{t,a}) \qquad \text{with} \qquad c^{\boldsymbol{\eta}}_{t,a} := (e_t)_\#\boldsymbol{\eta}_a,\ \boldsymbol{v}^{\boldsymbol{\eta}}_{t,a}c^{\boldsymbol{\eta}}_{t,a} := (e_t)_\#(\dot\omega(t)\boldsymbol{\eta}_a)$$

maps optimal solutions for the first problem into optimal solutions for the second one. Since under this correspondence (6.7) reduces to (6.6), the two pressure fields coincide.

The following crucial regularity result for the pressure field has been obtained in [5], improving in the time variable the regularity $\partial_{x_i}p \in \mathcal{M}_{\mathrm{loc}}\big((0,1)\times D\big)$ obtained by Brenier in [16].

Theorem 6.3 (Regularity of pressure) *Let $(c,\boldsymbol{v})$ be an optimal pair for the Eulerian-Lagrangian model, and let p be the pressure field identified by (6.7). Then $\partial_{x_i}p \in L^2_{\mathrm{loc}}((0,1);\mathcal{M}(D))$ and*

$$p \in L^2_{\mathrm{loc}}\big((0,1);BV_{\mathrm{loc}}(D)\big) \subset L^2_{\mathrm{loc}}\big((0,1);L^{d/(d-1)}_{\mathrm{loc}}(D)\big).$$

In the case $D = \mathbb{T}^d$ the same properties hold globally in space, i.e. replacing $BV_{\mathrm{loc}}(D)$ with $BV(\mathbb{T}^d)$ and $L^{d/(d-1)}_{\mathrm{loc}}(D)$ with $L^{d/(d-1)}(\mathbb{T}^d)$.

The $L^1_{\rm loc}$ integrability of p allows much stronger variations in the Lagrangian model, that give rise to possibly nonsmooth densities, which may even vanish.

From now one we shall confine our discussion to the case of the flat torus $\mathbb{T}^d$, as our arguments involve some *global* smoothing that becomes more technical, and needs to be carefully checked in more general situations. We also set $\mu_{\mathbb{T}} = \mu_{\mathbb{T}^d}$ and denote by $d_{\mathbb{T}}$ the Riemannian distance in $\mathbb{T}^d$ (i.e. the distance modulo 1 in $\mathbb{R}^d/\mathbb{Z}^d$). In the next theorem we consider generalized flows $\boldsymbol{\nu}$ with *bounded compression*, defined by the property $\rho^{\boldsymbol{\nu}} \in L^\infty\left((0,1)\times D\right)$.

Theorem 6.4 *Let $\boldsymbol{\eta}$ be an optimal incompressible flow in $\mathbb{T}^d$ between η and γ. Then*

$$\langle p, \rho^{\boldsymbol{\nu}} - 1\rangle \leq \mathscr{A}_1(\boldsymbol{\nu}) - \mathscr{A}_1(\boldsymbol{\eta}) \tag{6.8}$$

for any generalized flow with bounded compression $\boldsymbol{\nu}$ between η and γ such that

$$\rho^{\boldsymbol{\nu}}(t,\cdot) = 1 \text{ for } t \text{ sufficiently close to } 0,\ 1. \tag{6.9}$$

If $p \in L^1([0,1]\times\mathbb{T}^d)$, the condition (6.9) *is not required for the validity of* (6.8).

Proof. Let $J := \{\rho^{\boldsymbol{\nu}}(t,\cdot) \neq 1\} \Subset (0,1)$ and let us first assume that $\rho^{\boldsymbol{\nu}}$ is smooth. If $\|\rho^{\boldsymbol{\nu}} - 1\|_{C^1} \leq 1/2$, then the result follows by Theorem 6.2. If not, for $\varepsilon > 0$ small enough $(1-\varepsilon)\boldsymbol{\eta} + \varepsilon\boldsymbol{\nu}$ is a slightly compressible generalized flow in the sense of Definition 6.1. Thus, we have

$$\varepsilon\langle p, \rho^{\boldsymbol{\nu}} - 1\rangle = \langle p, \rho^{(1-\varepsilon)\boldsymbol{\eta}+\varepsilon\boldsymbol{\nu}} - 1\rangle \leq \mathscr{A}_1((1-\varepsilon)\boldsymbol{\eta} + \varepsilon\boldsymbol{\nu}) - \mathscr{A}_1(\boldsymbol{\eta}) = \varepsilon\left(\mathscr{A}_1(\boldsymbol{\nu}) - \mathscr{A}_1(\boldsymbol{\eta})\right),$$

and this proves the statement whenever $\rho^{\boldsymbol{\nu}}$ is smooth.

If $\rho^{\boldsymbol{\nu}}$ is not smooth, we need a regularization argument. Let us assume first that $\rho^{\boldsymbol{\nu}}$ is smooth in time, uniformly with respect to x, but not in space. We fix a cut-off function $\chi \in C^1_c(0,1)$ strictly positive on a neighbourhood of J and define, for $y \in \mathbb{R}^d$, the maps $T_{\varepsilon,y} : \tilde\Omega(\mathbb{T}^d) \to \tilde\Omega(\mathbb{T}^d)$ by

$$T_{\varepsilon,y}(\omega, a) := (\omega + \varepsilon y\chi, a), \qquad (\omega, a) \in \Omega(\mathbb{T}^d).$$

Then, we set $\boldsymbol{\nu}_\varepsilon := \int_{\mathbb{R}^d} (T_{\varepsilon,y})_\#\boldsymbol{\nu}\phi(y)\,dy$, where $\phi : \mathbb{R}^d \to [0,+\infty)$ is a standard convolution kernel. It is easy to check that $\boldsymbol{\nu}_\varepsilon$ still connects η to γ, and that

$$\rho^{\boldsymbol{\nu}_\varepsilon}(t,\cdot) = \rho^{\boldsymbol{\nu}}(t,\cdot) * \phi_{\varepsilon\chi(t)} \qquad \forall t \in [0,1],$$

where $\phi_\varepsilon(x) = \varepsilon^{-d}\phi(x/\varepsilon)$. Since

$$\lim_{\varepsilon\downarrow 0} \mathscr{A}_1(\boldsymbol{\nu}_\varepsilon) = \lim_{\varepsilon\downarrow 0} \int_{\mathbb{R}^d}\int_{\tilde\Omega(\mathbb{T}^d)}\int_0^1 |\dot\omega(t) + \varepsilon y\dot\chi(t)|^2\,dt\,d\boldsymbol{\nu}(\omega,a)\phi(y)\,dy = \mathscr{A}_1(\boldsymbol{\nu})$$

we can pass to the limit in (6.8) with $\boldsymbol{\nu}_\varepsilon$ in place of $\boldsymbol{\nu}$, which are smooth.

In the general case we fix a convolution kernel with compact support $\varphi(t)$ and, with the same choice of χ done before, we define the maps

$$T_\varepsilon(\omega, a)(t) := \left(\int_0^1 \omega(t - s\varepsilon\chi(t))\varphi(s)\,ds, a\right).$$

Setting $\boldsymbol{\nu}_\varepsilon = (T_\varepsilon)_\# \boldsymbol{\nu}$, it is easy to check that $\mathscr{A}_1(\boldsymbol{\nu}_\varepsilon) \to \mathscr{A}_1(\boldsymbol{\nu})$ and that

$$\rho^{\boldsymbol{\nu}_\varepsilon}(t,x) = \int_0^1 \rho^{\boldsymbol{\nu}}(t - s\varepsilon\chi(t), x)\varphi(s)\, ds$$

are smooth in time, uniformly in x. So, by applying (6.8) with $\boldsymbol{\nu}_\varepsilon$ in place of $\boldsymbol{\nu}$, we obtain the inequality in the limit.

Finally, if p is globally integrable, we can approximate any generalized flow with bounded compression $\boldsymbol{\nu}$ between η and γ by transforming ω into $\omega \circ \psi_\varepsilon$, where $\psi_\varepsilon : [0,1] \to [0,1]$ is defined by $\psi_\varepsilon(t) := \frac{1}{1-2\varepsilon} \int_0^t \chi_{[\varepsilon, 1-\varepsilon]}(s) ds$ (so that ψ_ε is constant for t close to 0 and 1). Passing to the limit as $\varepsilon \downarrow 0$ we obtain the inequality even without the condition $\rho^{\boldsymbol{\nu}}(t,\cdot) = 1$ for t close to 0, 1.

□

Remark 6.5 (Smoothing of flows and plans) Notice that the same smoothing argument can be used to prove this statement: given a flow $\boldsymbol{\eta}$ between $\eta = \eta_a \otimes \mu_{\mathbb{T}}$ and $\gamma = \gamma_a \otimes \mu_{\mathbb{T}}$ (not necessarily with bounded compression), we can find flows with bounded compression $\boldsymbol{\eta}^\varepsilon$ connecting $\eta^\varepsilon := (\eta_a) * \phi_\varepsilon \otimes \mu_{\mathbb{T}}$ to $\gamma^\varepsilon := (\gamma_a) * \phi_\varepsilon \otimes \mu_{\mathbb{T}}$, with $\mathscr{A}_T(\boldsymbol{\eta}^\varepsilon) = \mathscr{A}_T(\boldsymbol{\eta})$ and

$$\int_{\tilde\Omega(\mathbb{T}^d)} \int_0^1 r_\varepsilon(\tau,\omega)\, d\tau\, d\boldsymbol{\eta}(\omega,a) = \int_{\tilde\Omega(\mathbb{T}^d)} \int_0^1 r(\tau,\omega)\, d\tau\, d\boldsymbol{\eta}^\varepsilon(\omega,a) \qquad \forall r \in L^1\left([0,1]\times\mathbb{T}^d\right)$$

(where, as usual, $r_\varepsilon(t,x) = r(t,\cdot) * \phi_\varepsilon(x)$). In order to have these properties, it suffices to define

$$\boldsymbol{\eta}^\varepsilon := \int_{\mathbb{R}^d} (\sigma_{\varepsilon y})_\# \boldsymbol{\eta}\, \phi(y)\, dy,$$

where $\sigma_z(\omega,a) = (\omega + z, a)$. Notice also that the "mollified plans" η^ε, γ^ε converge to η, γ in $(\Gamma(\mathbb{T}^d), \bar\delta)$: if we consider the map $S^\varepsilon_y : \mathbb{T}^d \to \Omega(\mathbb{T}^d)$ given by $x \mapsto \omega_x(t) := x + \varepsilon t y$, the generalized incompressible flow $\boldsymbol{\nu}^\varepsilon = \boldsymbol{\nu}^\varepsilon_a \otimes \mu_{\mathbb{T}}$, with

$$\boldsymbol{\nu}^\varepsilon_a := \int_{\mathbb{R}^d} (S^\varepsilon_y)_\# \lambda_a\, \phi(y)\, dy,$$

connects in $[0,1]$ the plan $\lambda = \lambda_a \otimes \mu_{\mathbb{T}}$ to $\lambda^\varepsilon = (\lambda_a * \phi_\varepsilon) \otimes \mu_{\mathbb{T}}$, with an action equal to $\varepsilon^2 \int_{\mathbb{R}^d} |y|^2 \phi(y)\, dy$.

In order to state necessary and sufficient optimality conditions at the level of single fluid paths, we have to take into account that the pressure field is not pointwise defined, and to choose a particular representative in its equivalence class, modulo negligible sets in spacetime. Henceforth, we define

$$\bar p(t,x) := \liminf_{\varepsilon \downarrow 0} p_\varepsilon(t,x), \tag{6.10}$$

where, thinking of $p(t,\cdot)$ as a 1-periodic function in $\mathbb{R}^d$, p_ε is defined by

$$p_\varepsilon(t,x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} p(t, x + \varepsilon y) e^{-|y|^2/2}\, dy.$$

Notice that p_ε is smooth and still 1-periodic. The choice of the heat kernel here is convenient, because of the semigroup property $p_{\varepsilon+\varepsilon'} = (p_\varepsilon)_{\varepsilon'}$. Recall that $\bar p$ is a representative, because at any Lebesgue point x of $p(t,\cdot)$ the limit of $p_\varepsilon(t,x)$ exists, and coincides with $p(t,x)$.

In order to handle passages to limits, we need also uniform pointwise bounds on p_ε; therefore we define

$$Mf(x) := \sup_{\varepsilon>0}\, (2\pi)^{-d/2}\int_{\mathbb{R}^d} |f|(x+\varepsilon y)e^{-|y|^2/2}\,dy, \qquad f\in L^1(\mathbb{T}^d). \tag{6.11}$$

We will use the following facts: first,

$$Mf_\varepsilon = \sup_{\varepsilon'>0}|f_\varepsilon|_{\varepsilon'} \le \sup_{\varepsilon'>0}(|f|_\varepsilon)_{\varepsilon'} \le \sup_{r>0}|f|_r = Mf$$

because of the semigroup property; second, standard maximal inequalities imply $\|Mf\|_{L^p(\mathbb{T}^d)} \le c_p\|f\|_{L^p(\mathbb{T}^d)}$ for all $p>1$. Setting $Mp(t,x) := Mp(t,\cdot)(x)$, by Theorem 6.3 we infer that $Mp \in L^2_{\rm loc}\big((0,1), L^{d/d-1}(\mathbb{T}^d)\big)$, so that in particular $Mp\in L^1_{\rm loc}\big((0,1)\times\mathbb{T}^d\big)$. This is the integrability assumption on p that will play a role in the rest of this section.

Definition 6.6 (q-minimizing path) *Let $\omega\in H^1((0,1);D)$ with $Mq(\tau,\omega)\in L^1(0,1)$. We say that ω is a* q-minimizing path *if*

$$\int_0^1 \frac12|\dot\omega(\tau)|^2 - q(\tau,\omega)\,d\tau \le \int_0^1 \frac12|\dot\omega(\tau)+\dot\delta(\tau)|^2 - q(\tau,\omega+\delta)\,d\tau$$

for all $\delta\in H^1_0((0,1);D)$ with $Mq(\tau,\omega+\delta)\in L^1(0,1)$.
Analogously, we say that ω is a locally q-minimizing path *if*

$$\int_s^t \frac12|\dot\omega(\tau)|^2 - q(\tau,\omega)\,d\tau \le \int_s^t \frac12|\dot\omega(\tau)+\dot\delta(\tau)|^2 - q(\tau,\omega+\delta)\,d\tau \tag{6.12}$$

for all $[s,t]\subset(0,1)$ and all $\delta\in H^1_0((s,t);D)$ with $Mq(\tau,\omega+\delta)\in L^1(s,t)$.

Remark 6.7 We notice that, for incompressible flows $\boldsymbol\eta$, the L^1 (resp. $L^1_{\rm loc}$) integrability of $Mq(\tau,\omega)$ imposed on the curves ω (and on their perturbations $\omega+\delta$) is satisfied $\boldsymbol\eta$-a.e. if $Mq\in L^1\big((0,1)\times\mathbb{T}^d\big)$ (resp. $Mq\in L^1_{\rm loc}\big((0,1)\times\mathbb{T}^d\big)$); this can simply be obtained first noticing that the incompressibility of $\boldsymbol\eta$ and Fubini's theorem give

$$\int_{\tilde\Omega(\mathbb{T}^d)}\int_J f(\tau,\omega)\,d\tau\,d\boldsymbol\eta(\omega,a) = \int_J\int_{\mathbb{T}^d} f(\tau,x)\,d\mu_{\mathbb{T}^d}(x)\,d\tau$$

for all nonnegative Borel functions f and all intervals $J\subset(0,1)$, and then applying this identity to $f=Mq$.

Theorem 6.8 (First necessary condition) *Let $\boldsymbol\eta=\boldsymbol\eta_a\otimes\mu_{\mathbb{T}}$ be any optimal incompressible flow on $\mathbb{T}^d$. Then, $\boldsymbol\eta$ is concentrated on locally $\bar p$-minimizing paths, where $\bar p$ is the precise representative of the pressure field p, and on $\bar p$-minimizing paths if $Mp\in L^1([0,1]\times\mathbb{T}^d)$.*

Proof. With no loss of generality we identify $\mathbb{T}^d$ with $\mathbb{R}^d/\mathbb{Z}^d$. Let $\boldsymbol{\eta}$ be an optimal incompressible flow and $[s,t] \subset (0,1)$. We fix a nonnegative function $\chi \in C^1_c(0,1)$ with $\{\chi > 0\} = (s,t)$. Given $\delta \in H^1_0\left([s,t];\mathbb{T}^d\right)$, $y \in \mathbb{R}^d$ and a Borel set $E \subset \tilde{\Omega}(\mathbb{T}^d)$, we define $T_{\varepsilon,y} : \tilde{\Omega}(\mathbb{T}^d) \to \tilde{\Omega}(\mathbb{T}^d)$ by

$$T_{\varepsilon,y}(\omega,a) := \begin{cases} (\omega,a) & \text{if } \omega \notin E; \\ (\omega + \delta + \varepsilon y \chi, a) & \text{if } \omega \in E \end{cases}$$

(of course, the sum is understood modulo 1) and $\boldsymbol{\nu}_{\varepsilon,y} := (T_{\varepsilon,y})_\# \boldsymbol{\eta}$.

It is easy to see that $\boldsymbol{\nu}_{\varepsilon,y}$ is a flow with bounded compression, since for all times τ the curves $\omega(\tau)$ are either left unchanged, or translated by the constant $\delta(\tau) + \varepsilon y \chi(\tau)$, so that the density produced by $\boldsymbol{\nu}_{\varepsilon,y}$ is at most 2, and equal to 1 outside the interval $[s,t]$.

Therefore, by Theorem 6.4 we get

$$\int_{\mathbb{T}^d} \int_s^t \bar{p}(\rho^{\boldsymbol{\nu}_{\varepsilon,y}} - 1)\, d\tau\, d\mu_{\mathbb{T}} \le \int_E \mathscr{A}_1(\omega + \delta + \varepsilon y \chi) - \mathscr{A}_1(\omega)\, d\boldsymbol{\eta}(\omega,a).$$

Rearranging terms, we get

$$\int_E \int_s^t \frac{1}{2}|\dot{\omega}|^2 - \bar{p}(\tau,\omega)\, d\tau\, d\boldsymbol{\eta}(\omega,a) \le \int_E \left[\int_s^t \frac{1}{2}|\dot{\omega} + \dot{\delta} + \varepsilon y \dot{\chi}|^2 - \bar{p}(\tau, \omega + \delta + \varepsilon y \chi)\, d\tau\right] d\boldsymbol{\eta}(\omega,a).$$

We can now average the above inequality using the heat kernel $\phi(y) = (2\pi)^{-d/2} e^{-|y|^2/2}$, and we obtain

$$\begin{aligned} &\int_E \int_s^t \frac{1}{2}|\dot{\omega}|^2 - \bar{p}(\tau,\omega)\, d\tau\, d\boldsymbol{\eta}(\omega,a) \\ &\qquad \le \int_E \left[\int_{\mathbb{R}^d} \int_s^t \frac{1}{2}|\dot{\omega} + \dot{\delta} + \varepsilon y \dot{\chi}|^2\, d\tau \phi(y)\, dy - \int_s^t p_{\varepsilon\chi(\tau)}(\tau, \omega + \delta)\, d\tau\right] d\boldsymbol{\eta}(\omega,a). \end{aligned}$$

Now, let $\mathcal{D} \subset H^1_0\left([s,t];\mathbb{T}^d\right)$ be a countable dense subset; by the arbitrariness of E and Remark 6.7 we infer the existence of a $\boldsymbol{\eta}$-negligible Borel set $B \subset \tilde{\Omega}(\mathbb{T}^d)$ such that $Mp(\tau,\omega) \in L^1(s,t)$ and

$$\int_s^t \frac{1}{2}|\dot{\omega}|^2 - \bar{p}(\tau,\omega)\, d\tau \le \int_{\mathbb{R}^d} \int_s^t \frac{1}{2}|\dot{\omega} + \dot{\delta} + \varepsilon y \dot{\chi}|^2\, d\tau \phi(y)\, dy - \int_s^t p_{\varepsilon\chi(\tau)}(\tau, \omega + \delta)\, d\tau$$

holds for all $\varepsilon = 1/n$, $\delta \in \mathcal{D}$ and $(\omega,a) \in \tilde{\Omega}(\mathbb{T}^d) \setminus B$. By a density argument, we see that the same inequality holds for all $\varepsilon = 1/n$, $\delta \in H^1_0\left([s,t];\mathbb{T}^d\right)$, and $\omega \in \tilde{\Omega}(\mathbb{T}^d) \setminus B$.

Now, if $Mp(\tau, \omega + \delta) \in L^1(s,t)$, we can use the bound $|p_\varepsilon| \le Mp$ to pass to the limit as $\varepsilon \downarrow 0$ to obtain that (6.12) holds with $q = \bar{p}$.

The proof of the global minimality property in the case when $p \in L^1([0,T] \times \mathbb{T}^d)$ is similar, just letting δ vary in $H^1_0\left([0,1];\mathbb{T}^d\right)$ and using a fixed function $\chi \in C^1([0,1])$ with $\chi(0) = \chi(1) = 0$ and $\chi > 0$ in $(0,1)$. □

In order to state the second necessary optimality condition fulfilled by minimizers, we need some preliminary definition. Let $q \in L^1([s,t] \times D)$ and let us define the cost $c_q^{s,t} : D \times D \to \overline{\mathbb{R}}$ of the minimal connection in $[s,t]$ between x and y, namely

$$c_q^{s,t}(x,y) := \inf \left\{ \int_s^t \frac{1}{2}|\dot{\omega}(\tau)|^2 - q(\tau,\omega)\,d\tau :\ \omega(s) = x,\ \omega(t) = y,\ Mq(\tau,\omega) \in L^1(s,t) \right\}, \quad (6.13)$$

with the convention $c_q^{s,t}(x,y) = +\infty$ if no admissible curve ω exists. Using this cost function $c_q^{s,t}$, we can consider the induced optimal transport problem, namely

$$W_{c_q^{s,t}}(\mu_1,\mu_2) := \inf \left\{ \int_{D\times D} c_q^{s,t}(x,y)\,d\lambda(x,y) :\ \lambda \in \Gamma(\mu_1,\mu_2),\ (c_q^{s,t})^+ \in L^1(\lambda) \right\}, \quad (6.14)$$

where $\Gamma(\mu_1,\mu_2)$ is the family of all probability measures λ in $D \times D$ whose first and second marginals are respectively μ_1 and μ_2. Again, we set by convention $W_{c_q^{s,t}}(\mu_1,\mu_2) = +\infty$ if no admissible λ exists.

Unlike most classical situations (see [26]), existence of an optimal λ is not guaranteed because $c_q^{s,t}$ are not lower semicontinuous in $D \times D$, and also it seems difficult to get lower bounds on $c_q^{s,t}$. It will be useful, however, the following *upper* bound on $W_{c_q^{s,t}}$:

Lemma 6.9 *If $Mq \in L^1([s,t] \times \mathbb{T}^d)$ there exists a nonnegative $\mu_{\mathbb{T}}$-integrable function $K_q^{s,t}$ satisfying*

$$c_q^{s,t}(x,y) \le K_q^{s,t}(x) + K_q^{s,t}(y) \qquad \forall x, y \in \mathbb{T}^d. \quad (6.15)$$

Remark 6.10 By (6.15) we deduce that, if $K_q^{s,t} \in L^1(\mu_1 + \mu_2)$, then $(c_q^{s,t})^+ \in L^1(\lambda)$ for all $\lambda \in \Gamma(\mu_1,\mu_2)$ and we have

$$\int_{\mathbb{T}^d\times\mathbb{T}^d} c_q^{s,t}(x,y)\,d\lambda(x,y) \le \int_{\mathbb{T}^d} K_q^{s,t}(w)\,d(\mu_1+\mu_2)(w) \qquad \forall \lambda \in \Gamma(\mu_1,\mu_2).$$

In particular, $W_{c_q^{s,t}}(\mu_1,\mu_2)$ as defined in (6.14) is not equal to $+\infty$.

Proof. Assume $s = 0$ and let $l = t/2$. Let us fix x, $y \in \mathbb{T}^d$; given $z \in \mathbb{T}^d$ we consider the projection on $\mathbb{T}^d$ of the Euclidean path

$$\omega_z(\tau) := \begin{cases} x + \frac{\tau}{l}(z-x) & \text{if } \tau \in [0,l]; \\ z + \frac{\tau-l}{l}(y-z) & \text{if } \tau \in [l,t]. \end{cases}$$

This path leads to the estimate

$$c_q^{0,t}(x,y) \le \frac{d_{\mathbb{T}}^2(x,z) + d_{\mathbb{T}}^2(z,y)}{2l} + \int_0^l Mq(\tau, x + \frac{\tau}{l}(z-x))\,d\tau + \int_l^t Mq(\tau, z + \frac{\tau-l}{l}(y-z))\,d\tau.$$

By integrating the free variable z with respect to $\mu_{\mathbb{T}}$, since $d_{\mathbb{T}} \le \frac{\sqrt{d}}{2}$ on $\mathbb{T}^d \times \mathbb{T}^d$, we get

$$c_q^{0,t}(x,y) \le \frac{d}{4l} + \int_{\mathbb{T}^d}\int_0^l Mq(\tau, x + \frac{\tau}{l}(z-x)) + Mq(l+\tau, z + \frac{\tau}{l}(y-z))\,d\tau\,d\mu_{\mathbb{T}}(z).$$

Therefore, the function

$$K_q^{0,t}(w) := \frac{d}{4l} + \int_{\mathbb{T}^d}\int_0^l Mq(\tau, w + \frac{\tau}{l}(z-w)) + Mq(l+\tau, z + \frac{\tau}{l}(w-z))\,d\tau\,d\mu_{\mathbb{T}}(z) \qquad (6.16)$$

fulfils (6.15). It is easy to check, using Fubini's theorem, that $K_q^{0,t}$ is $\mu_{\mathbb{T}}$-integrable in $\mathbb{T}^d$. Indeed,

$$\begin{aligned}
\int_{\mathbb{T}^d} K_q^{0,t}(w)\,d\mu_{\mathbb{T}}(w) &= \frac{d}{4l} + \int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\int_0^l Mq(\tau, w + \frac{\tau}{l}(z-w))\,d\tau\,d\mu_{\mathbb{T}}(z)\,d\mu_{\mathbb{T}}(w) \\
&+ \int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\int_0^l Mq(l+\tau, z + \frac{\tau}{l}(w-z))\,d\tau\,d\mu_{\mathbb{T}}(w)\,d\mu_{\mathbb{T}}(z) \\
&= \frac{d}{4l} + \int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\int_0^l Mq(\tau, w + \frac{\tau}{l}y)\,d\tau\,d\mu_{\mathbb{T}}(y)\,d\mu_{\mathbb{T}}(w) \\
&+ \int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\int_0^l Mq(l+\tau, z + \frac{\tau}{l}y)\,d\tau\,d\mu_{\mathbb{T}}(z)\,d\mu_{\mathbb{T}}(y) \\
&= \frac{d}{4l} + \int_0^l\int_{\mathbb{T}^d}\int_{\mathbb{T}^d} Mq(\tau, w + \frac{\tau}{l}y) + Mq(l+\tau, w + \frac{\tau}{l}y)\,d\mu_{\mathbb{T}}(w)\,d\mu_{\mathbb{T}}(y)\,d\tau \\
&= \frac{d}{4l} + \int_0^t\int_{\mathbb{T}^d} Mq(\tau, w)\,d\mu_{\mathbb{T}}(w)\,d\tau < +\infty.
\end{aligned}$$

□

In the proof of the next theorem we are going to use the *measurable selection theorem* (see [19, Theorems III.22 and III.23]): if $(A, \mathcal{A}, \nu)$ is a measure space, X is a Polish space and $E \subset A \times X$ is $\mathcal{A}_\nu \otimes \mathcal{B}(X)$-measurable, where $\mathcal{A}_\nu$ is the ν-completion of $\mathcal{A}$, then:

(i) the projection $\pi_A(E)$ of E on A is $\mathcal{A}_\nu$-measurable;

(ii) there exists a $(\mathcal{A}_\nu, \mathcal{B}(X))$-measurable map $\sigma : \pi(E) \to X$ such that $(x, \sigma(x)) \in E$ for ν-a.e. $x \in \pi_A(E)$.

The next theorem will provide a new necessary optimality condition involving not only the path that should be followed between x and y (which, as we proved, should minimize the Lagrangian $\mathcal{L}_{\bar{p}}$ in (1.8)), but also the "weights" given to the paths. We observe that, when a variation of these weights is performed, new flows $\tilde{\boldsymbol{\eta}}$ between η and γ are built which need not be of bounded compression, for which $(e_t)_\#\tilde{\boldsymbol{\eta}}$ might be even singular with respect to $\mu_{\mathbb{T}}$; therefore we can't use directly them in the variational principle (6.8); however, this difficulty can be overcome by the smoothing procedure in Remark 6.5.

Theorem 6.11 (Second necessary condition) *Let $\boldsymbol{\eta} = \boldsymbol{\eta}_a \otimes \mu_{\mathbb{T}}$ be an optimal incompressible flow on $\mathbb{T}^d$ between η and γ. Then, for all intervals $[s,t] \subset (0,1)$, $W_{c_{\bar{p}}^{s,t}}(\eta_a, \gamma_a) \in \mathbb{R}$ and the plan $(e_s, e_t)_\#\boldsymbol{\eta}_a$ is optimal, relative to the cost $c_{\bar{p}}^{s,t}$ defined in* (6.13), *for $\mu_{\mathbb{T}}$-a.e. a.*

Proof. Let $[s,t]\subset(0,1)$ be fixed. Since

$$\begin{aligned}\int_{\mathbb{T}^d}\int_{\mathbb{T}^d\times\mathbb{T}^d} c_{\bar p}^{s,t}(x,y)\,d(e_s,e_t)_\#\boldsymbol{\eta}_a\,d\mu_{\mathbb{T}}(a) &\le \int_{\mathbb{T}^d}\int_{\Omega(\mathbb{T}^d)}\int_s^t \frac{1}{2}|\dot\omega(\tau)|^2-\bar p(\tau,\omega)\,d\tau d\boldsymbol{\eta}_a(\omega)\,d\mu_{\mathbb{T}}(a)\\ &= (t-s)\overline{\delta}^2(\eta,\gamma), \end{aligned}\tag{6.17}$$

it suffices to show that

$$(t-s)\overline{\delta}^2(\eta,\gamma)\le\int_{\mathbb{T}^d} W_{c_{\bar p}^{s,t}}(\eta_a^s,\gamma_a^t)\,d\mu_{\mathbb{T}}(a).\tag{6.18}$$

We are going to prove this fact by a smoothing argument. We set $\eta^s=\eta_a^s\otimes\mu_{\mathbb{T}}$, $\gamma^t=\gamma_a^t\otimes\mu_{\mathbb{T}}$, with $\eta_a^s=(e_s)_\#\boldsymbol{\eta}_a$, $\gamma_a^t=(e_t)_\#\boldsymbol{\eta}_a$. Recall that Remark 3.1 gives

$$\overline{\delta}(\eta,\eta^s)=s\overline{\delta}(\eta,\gamma),\qquad \overline{\delta}(\gamma^t,\gamma)=(1-t)\overline{\delta}(\eta,\gamma).$$

First, we notice that Lemma 6.9 gives

$$\begin{aligned}\int_{\mathbb{T}^d} W_{c_{-|\bar p|}^{s,t}}(\eta_a^s,\gamma_a^t)\,d\mu_{\mathbb{T}}(a) &\le \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} K_{-|\bar p|}^{s,t}(w)\,d(\eta_a^s+\gamma_a^t)(w)\,d\mu_{\mathbb{T}}(a)\\ &=2\int_{\mathbb{T}^d} K_{-|\bar p|}^{s,t}(w)\,d\mu_{\mathbb{T}}(w)<+\infty.\end{aligned}\tag{6.19}$$

We also remark that, since $\tau\mapsto\|p_\varepsilon(\tau,\cdot)\|_\infty$ is integrable in (s,t), for any $\varepsilon>0$ the cost $c_{p_\varepsilon}^{s,t}$ is bounded both from above and below. Next, we show that

$$c_{\bar p}^{s,t}(x,y)\ge\limsup_{\varepsilon\downarrow 0} c_{p_\varepsilon}^{s,t}(x,y)\qquad\forall (x,y)\in\mathbb{T}^d\times\mathbb{T}^d.\tag{6.20}$$

Indeed, let $\omega\in H^1\left([s,t];\mathbb{T}^d\right)$ with $\omega(s)=x$, $\omega(t)=y$ and $Mp(\tau,\omega)\in L^1(s,t)$ (if there is no such ω, there is nothing to prove). By the pointwise bound $|p_\varepsilon|\le Mp$ and Lebesgue's theorem, we get

$$\int_s^t\frac{1}{2}|\dot\omega(\tau)|^2-\bar p(\tau,\omega)\,d\tau=\lim_{\varepsilon\downarrow 0}\int_s^t\frac{1}{2}|\dot\omega(\tau)|^2-p_\varepsilon(\tau,\omega)\,d\tau.$$

By the $L^1(L^\infty)$ bound on Mp_ε, the curve ω is admissible also for the variational problem defining $c_{p_\varepsilon}^{s,t}$, therefore the above limit provides an upper bound on $\limsup_\varepsilon c_{p_\varepsilon}^{s,t}(x,y)$. By minimizing with respect to ω we obtain (6.20).

By (6.19) and the pointwise bound $\bar p\ge -|\bar p|$ we infer that the positive part of $W_{c_{\bar p}^{s,t}}(\eta_a^s,\gamma_a^t)$ is $\mu_{\mathbb{T}}$-integrable. Let now $\delta>0$ be fixed, and let us consider the compact space $X:=\mathscr{P}(\mathbb{T}^d\times\mathbb{T}^d)$ and the $\mathcal{B}(\mathbb{T}^d)_{\mu_{\mathbb{T}}}\otimes\mathcal{B}(X)$-measurable set

$$E:=\left\{(a,\lambda)\in\mathbb{T}^d\times X:\ \lambda\in\Gamma(\eta_a^s,\gamma_a^t),\ \int_{\mathbb{T}^d\times\mathbb{T}^d} c_{\bar p}^{s,t}(x,y)\,d\lambda<\delta+\left(W_{c_{\bar p}^{s,t}}(\eta_a^s,\gamma_a^t)\vee-\frac{1}{\delta}\right)\right\}$$

(we skip the proof of the measurability, that is based on tedious but routine arguments). Since $W_{c_{\bar p}^{s,t}}(\eta_a^s,\gamma_a^t)<+\infty$ for $\mu_{\mathbb{T}}$-a.e. a, we obtain that for $\mu_{\mathbb{T}}$-a.e. $a\in\mathbb{T}^d$ there exists $\lambda\in\Gamma(\eta_a^s,\gamma_a^t)$

with $(a,\lambda)\in E$. Thanks to the measurable selection theorem we can select a Borel family $a\mapsto\lambda_a\in\mathscr{P}(\mathbb{T}^d\times\mathbb{T}^d)$ such that $\lambda_a\in\Gamma(\eta^s_a,\gamma^t_a)$ and

$$\int_{\mathbb{T}^d\times\mathbb{T}^d} c^{s,t}_{\bar p}(x,y)\,d\lambda_a<\delta+\Big(W_{c^{s,t}_{\bar p}}(\eta^s_a,\gamma^t_a)\vee-\frac{1}{\delta}\Big)\qquad\text{for }\mu_{\mathbb{T}}\text{-a.e. }a\in\mathbb{T}^d.$$

By Lemma 6.9 and Remark 6.10 we get

$$c^{s,t}_{p_\varepsilon}(x,y)\le K^{s,t}_{p_\varepsilon}(x)+K^{s,t}_{p_\varepsilon}(y)\le K^{s,t}_{p}(x)+K^{s,t}_{p}(y)\qquad\forall x,y\in\mathbb{T}^d$$

and

$$\int_{\mathbb{T}^d}\int_{\mathbb{T}^d\times\mathbb{T}^d}K^{s,t}_p(x)+K^{s,t}_p(y)\,d\lambda_a\,d\mu_{\mathbb{T}}(a)=\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}K^{s,t}_p\,d(\eta^s_a+\gamma^t_a)\,d\mu_{\mathbb{T}}(a)<+\infty$$

(we used the pointwise bound $Mp_\varepsilon\le Mp$ and the fact that $q\mapsto K^{s,t}_q$ has a monotone dependence upon Mq, see (6.16)). Therefore (6.20) and Fatou's lemma give

$$\delta+\int_{\mathbb{T}^d}\Big(W_{c^{s,t}_{\bar p}}(\eta^s_a,\gamma^t_a)\vee-\frac{1}{\delta}\Big)\,d\mu_{\mathbb{T}}(a)\ge\limsup_{\varepsilon\downarrow0}\int_{\mathbb{T}^d}\int_{\mathbb{T}^d\times\mathbb{T}^d}c^{s,t}_{p_\varepsilon}(x,y)d\lambda_a\,d\mu_{\mathbb{T}}(a).\tag{6.21}$$

Still thanks to the measurable selection theorem, we can find a Borel map $(x,y,a)\mapsto\omega^{x,y}_{a,\varepsilon}\in C\left([s,t];\mathbb{T}^d\right)$ with $\omega^{x,y}_{a,\varepsilon}(s)=x$, $\omega^{x,y}_{a,\varepsilon}(t)=y$, $Mp_\varepsilon(\tau,\omega^{x,y}_{a,\varepsilon})\in L^1(s,t)$ and

$$\int_s^t\frac12|\dot\omega^{x,y}_{a,\varepsilon}|^2-p_\varepsilon(\tau,\omega^{x,y}_{a,\varepsilon})\,d\tau<\delta+c^{s,t}_{p_\varepsilon}(x,y)\qquad\text{for }\lambda_a\otimes\mu_{\mathbb{T}}\text{-a.e. }(x,y,a).$$

Let $\boldsymbol{\lambda}^\varepsilon=\boldsymbol{\lambda}^\varepsilon_a\otimes\mu_{\mathbb{T}}$ be the push-forward, under the map $(x,y,a)\mapsto\omega^{x,y}_{a,\varepsilon}$, of the measure $\lambda_a\otimes\mu_{\mathbb{T}}$; by construction this measure fulfils $(e_s)_\#\boldsymbol{\lambda}^\varepsilon_a=\eta^s_a$, $(e_t)_\#\boldsymbol{\lambda}^\varepsilon_a=\gamma^t_a$, (because the marginals of λ_a are η^s_a and γ^t_a), therefore it connects η^s to γ^t in $[s,t]$. Then, from (6.21) we get

$$2\delta+\int_{\mathbb{T}^d}\Big(W_{c^{s,t}_{\bar p}}(\eta^s_a,\gamma^t_a)\vee-\frac{1}{\delta}\Big)\,d\mu_{\mathbb{T}}(a)\ge\limsup_{\varepsilon\downarrow0}\int_{C([s,t];\mathbb{T}^d)\times\mathbb{T}^d}\int_s^t\frac12|\dot\omega(\tau)|^2-p_\varepsilon(\tau,\omega)\,d\tau\,d\boldsymbol{\lambda}^\varepsilon(\omega,a).$$

Eventually, Remark 6.5 provides us with a flow with bounded compression $\hat{\boldsymbol{\lambda}}^\varepsilon$ connecting $\eta^{s,\varepsilon}$ to $\gamma^{t,\varepsilon}$ in $[s,t]$ with

$$2\delta+\int_{\mathbb{T}^d}\Big(W_{c^{s,t}_{\bar p}}(\eta_a,\gamma_a)\vee-\frac{1}{\delta}\Big)\,d\mu_{\mathbb{T}}(a)\ge\limsup_{\varepsilon\downarrow0}\int_{C([s,t];\mathbb{T}^d)\times\mathbb{T}^d}\int_s^t\frac12|\dot\omega(\tau)|^2-\bar p(\tau,\omega)\,d\tau\,d\hat{\boldsymbol{\lambda}}^\varepsilon(\omega,a).\tag{6.22}$$

Since $\eta^{s,\varepsilon}\to\eta^s$ and $\gamma^{t,\varepsilon}\to\gamma^t$ in $(\Gamma(\mathbb{T}^d),\overline{\delta})$, we can find (by scaling $\boldsymbol{\eta}$ from $[0,s]$ to $[0,s_\varepsilon]$ and from $[t,1]$ to $[t_\varepsilon,1]$, and using repeatedly the concatenation, see Remark 3.2) generalized flows $\boldsymbol{\nu}^\varepsilon$ between γ and η in $[0,1]$, $s_\varepsilon\uparrow s$, $t_\varepsilon\downarrow t$ satisfying:

(a) $\boldsymbol{\nu}^\varepsilon$ connects η to η^s in $[0,s_\varepsilon]$, η^s to $\eta^{s,\varepsilon}$ in $[s_\varepsilon,s]$, $\gamma^{t,\varepsilon}$ to γ^t in $[t,t_\varepsilon]$, γ^t to γ in $[t_\varepsilon,1]$ and is incompressible in all these time intervals;

(b) the restriction of $\boldsymbol{\nu}^\varepsilon$ to $[s,t]$ coincides with $\hat{\boldsymbol{\lambda}}^\varepsilon$;

(c) the action of $\boldsymbol{\nu}^\varepsilon$ in $[0,s]$ converges to $\overline{\delta}^2(\eta,\eta^s) = s^2\overline{\delta}^2(\eta,\gamma)$, and the action of $\boldsymbol{\nu}^\varepsilon$ in $[t,1]$ converges to $\overline{\delta}^2(\gamma^t,\gamma) = (1-t)^2\overline{\delta}^2(\eta,\gamma)$.

Since $\boldsymbol{\nu}^\varepsilon$ is a flow with bounded compression connecting η to γ we use (6.8) and the incompressibility in $[0,1]\setminus[s,t]$ to obtain

$$\int_{\tilde{\Omega}(\mathbb{T}^d)}\int_0^1 \frac{1}{2}|\dot{\omega}(\tau)|^2\,d\tau\,d\boldsymbol{\nu}^\varepsilon(\omega,a) - \int_{\tilde{\Omega}(\mathbb{T}^d)}\int_s^t \bar{p}(\tau,\omega)\,d\tau\,d\boldsymbol{\nu}^\varepsilon(\omega,a) \geq \overline{\delta}^2(\eta,\gamma) \tag{6.23}$$

for all $\varepsilon > 0$. Taking into account that (b) and (c) imply

$$\int_{\tilde{\Omega}(\mathbb{T}^d)}\int_0^s \frac{1}{2}|\dot{\omega}(\tau)|^2\,d\tau\,d\boldsymbol{\nu}^\varepsilon(\omega,a) \to s\overline{\delta}^2(\eta,\gamma)$$

and

$$\int_{\tilde{\Omega}(\mathbb{T}^d)}\int_t^1 \frac{1}{2}|\dot{\omega}(\tau)|^2\,d\tau\,d\boldsymbol{\nu}^\varepsilon(\omega,a) \to (1-t)\overline{\delta}^2(\eta,\gamma),$$

from (6.22) and (6.23) we get

$$2\delta + \int_{\mathbb{T}^d}\Big(W_{c_{\bar{p}}^{s,t}}(\eta_a^s,\gamma_a^t) \vee -\frac{1}{\delta}\Big)\,d\mu_{\mathbb{T}}(a) \geq (1-s-(1-t))\overline{\delta}^2(\eta,\gamma) = (t-s)\overline{\delta}^2(\eta,\gamma). \tag{6.24}$$

Letting $\delta \downarrow 0$ we obtain the $\mu_{\mathbb{T}}$-integrability of $W_{c_{\bar{p}}^{s,t}}(\eta_a^s,\gamma_a^t)$ and (6.18). $\square$

A byproduct of the above proof is that equalities hold in (6.17), (6.18), and therefore

$$\begin{aligned}&\int_{\mathbb{T}^d}\int_{\Omega(\mathbb{T}^d)}\Big(\int_s^t \frac{1}{2}|\dot{\omega}(\tau)|^2 - \bar{p}(\tau,\omega)\,d\tau - c_{\bar{p}}^{s,t}(\omega(s),\omega(t))\Big)\,d\boldsymbol{\eta}_a(\omega)\,d\mu_{\mathbb{T}}(a) \qquad (6.25)\\ &= \int_{\mathbb{T}^d}\int_{\Omega(\mathbb{T}^d)}\int_s^t \frac{1}{2}|\dot{\omega}(\tau)|^2 - \bar{p}(\tau,\omega)\,d\tau\,d\boldsymbol{\eta}_a(\omega)\,d\mu_{\mathbb{T}}(a) - \int_{\mathbb{T}^d} W_{c_{\bar{p}}^{s,t}}(\eta_a^s,\gamma_a^t)\,d\mu_{\mathbb{T}}(a) = 0.\end{aligned}$$

This yields in particular also the first optimality condition. However, as the proof of Theorem 6.11 is much more technical than the one presented in Theorem 6.8, we decided to present both.

Now we show that the optimality conditions in Theorems 6.8 and 6.11 are also sufficient, even in the case of a general compact manifold without boundary D.

Theorem 6.12 (Sufficient condition) *Assume that $\boldsymbol{\eta} = \boldsymbol{\eta}_a \otimes \mu$ is a generalized incompressible flow in D between η and γ, and assume that for some map q the following properties hold:*

(a) $Mq \in L^1((0,1)\times D)$ and $\boldsymbol{\eta}$ is concentrated on q-minimizing paths;

(b) the plan $(e_0,e_1)_\#\boldsymbol{\eta}_a$ is optimal, relative to the cost $c_q^{0,1}$ defined in (6.13), for μ_D-a.e. a.

Then $\boldsymbol{\eta}$ is optimal and q is the pressure field. In addition, if (a), (b) are replaced by

(a') $Mq \in L^1_{\rm loc}((0,1)\times D)$ and $\boldsymbol{\eta}$ is concentrated on locally q-minimizing paths;

(b') for all intervals $[s,t]\subset(0,1)$, the plan $(e_s,e_t)_\#\boldsymbol{\eta}_a$ is optimal, relative to the cost $c_q^{s,t}$ defined in (6.13), for μ_D-a.e. a,

the same conclusions hold.

Proof. Assume first that (a) and (b) hold, and assume without loss of generality that $\int_D q(t,\cdot)\,d\mu_D = 0$ for almost all $t\in(0,1)$. Recalling that, thanks to the global integrability of Mq, any generalized incompressible flow $\boldsymbol{\nu}=\boldsymbol{\nu}_a\otimes\mu_D$ between η and γ is concentrated on curves ω with $Mq(\tau,\omega)\in L^1(0,1)$ (see Remark 6.7), we have

$$\begin{aligned}\mathscr{A}_1(\boldsymbol{\nu}) &= \int_D\int_{\Omega(D)}\int_0^1 \frac{1}{2}|\dot\omega|^2 - q(\tau,\omega)\,d\tau\,d\boldsymbol{\nu}_a(\omega)\,d\mu_D(a) \\ &\geq \int_D\int_{D\times D} c_q^{0,1}(x,y)\,d(e_0,e_1)_\#\boldsymbol{\nu}_a\,d\mu_D(a) \geq \int_D W_{c_q^{0,1}}(\eta_a,\gamma_a)\,d\mu_D(a).\end{aligned} \tag{6.26}$$

When $\boldsymbol{\nu}=\boldsymbol{\eta}$ the first inequality is an equality, because $\boldsymbol{\eta}$ is concentrated on q-minimizing paths, as well as the second inequality, because of the optimality of the plan $(e_0,e_1)_\#\boldsymbol{\eta}_a$. This proves that $\boldsymbol{\eta}$ is optimal. Moreover, by using the inequality in (6.26) with a flow $\boldsymbol{\nu}$ with bounded compression, one obtains

$$\mathscr{A}_1(\boldsymbol{\nu}) \geq \mathscr{A}_1(\boldsymbol{\eta}) + \langle q, \rho^{\boldsymbol{\nu}}-1\rangle.$$

Considering almost incompressible flows $\boldsymbol{\nu}$ arising by a smooth perturbation of $\boldsymbol{\eta}$ as described at the beginning of this section (see (6.1) in particular), the same argument used to obtain (6.6) gives that q satisfies (6.6), so that q is the pressure field.

In the case when (a)' and (b)' hold, by localizing in all intervals $[s,t]\subset(0,1)$ the previous argument (see Remark 3.2), one obtains that

$$(t-s)\int_{\tilde\Omega(D)}\int_s^t \frac{1}{2}|\dot\omega|^2\,d\tau\,d\boldsymbol{\eta}(\omega,a) = \bar\delta^2(\gamma_s,\gamma_t),$$

where $\gamma_s=(e_s,\pi_D)_\#\boldsymbol{\eta}$ and $\gamma_t=(e_t,\pi_D)_\#\boldsymbol{\eta}$. Letting $s\downarrow 0$ and $t\uparrow 1$ we obtain the optimality of $\boldsymbol{\eta}$. □

A byproduct of the previous result is a new variational principle satisfied, at least locally in time, by the pressure field. Up to a restriction to a smaller time interval we shall assume that $Mp\in L^1\big([0,1]\times\mathbb{T}^d\big)$.

Corollary 6.13 (Variational characterization of the pressure) *Let η, $\gamma\in\Gamma(\mathbb{T}^d)$ and let p be the unique pressure field induced by the constant speed geodesics in $[0,1]$ between $\eta=\eta_a\otimes\mu_{\mathbb{T}}$ and $\gamma=\gamma_a\otimes\mu_{\mathbb{T}}$. Assume that $Mp\in L^1\big([0,1]\times\mathbb{T}^d\big)$ and, with no loss of generality, $\int_{\mathbb{T}^d} p(t,\cdot)\,d\mu_{\mathbb{T}}=0$. Then $\bar p$ maximizes the functional*

$$q\mapsto\Psi(q) := \int_{\mathbb{T}^d} W_{c_q^{0,1}}(\eta_a,\gamma_a)\,d\mu_{\mathbb{T}}(a) + \int_0^1\int_{\mathbb{T}^d} q(\tau,x)\,d\mu_{\mathbb{T}}(x)\,d\tau$$

among all functions $q:[0,1]\times\mathbb{T}^d\to\mathbb{R}$ with $Mq\in L^1([0,1]\times\mathbb{T}^d)$.

Proof. We first remark that the functional Ψ is invariant under sum of functions depending on t only, so we can assume that the spatial means of any function q vanish.
From (6.25) we obtain that

$$\int_{\mathbb{T}^d} W_{c^{0,1}_{\bar p}}(\eta_a,\gamma_a)\,d\mu_{\mathbb{T}}(a) = \int_{\mathbb{T}^d}\int_{\Omega(\mathbb{T}^d)}\int_0^1 \frac{1}{2}|\dot\omega(\tau)|^2 - \bar p(\tau,\omega)\,d\tau\,d\boldsymbol{\eta}_a(\omega)\,d\mu_{\mathbb{T}}(a).$$

By the incompressibility constraint, in the right hand side $\bar p$ can be replaced by any function q whose spatial means vanish and, if $Mq \in L^1\big([0,1]\times\mathbb{T}^d\big)$, the resulting integral bounds from above $\int_{\mathbb{T}^d} W_{c^{0,1}_q}(\eta_a,\gamma_a)\,d\mu_{\mathbb{T}}(a)$, as we proved in (6.26). □

References

[1] L.Ambrosio, N.Fusco & D.Pallara: *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, 2000.

[2] L.Ambrosio: *Transport equation and Cauchy problem for BV vector fields.* Inventiones Mathematicae, **158** (2004), 227–260.

[3] L.Ambrosio, N.Gigli & G.Savaré: *Gradient flows in metric spaces and in the Wasserstein space of probability measures.* Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.

[4] L.Ambrosio, S.Lisini & G.Savaré: *Stability of flows associated to gradient vector fields and convergence of iterated transport maps.* Manuscripta Mathematica, **121** (2006), 1–50.

[5] L.Ambrosio & A.Figalli: *On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations.* Calc. Var. Partial Differential Equations, **31** (2008), no. 4, 497-509.

[6] V.Arnold: *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits.* (French) Ann. Inst. Fourier (Grenoble), **16** (1966), fasc. 1, 319–361.

[7] V.Bangert: *Minimal measures and minimizing closed normal one-currents.* Geom. Funct. Anal., **9** (1999), 413–427.

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

[9] P.Bernard & B.Buffoni: *Optimal mass transportation and Mather theory.* J. Eur. Math. Soc., **9** (2007), no. 1, 85-121.

[10] P.Bernard: *Young measures, superposition, and transport.* Indiana Univ. Math. J., **57** (2008), no. 1, 247-275.

[11] M.Bernot, V.Caselles & J.M.Morel: *Traffic plans.* Publ. Mat., **49** (2005), no. 2, 417–451.

[12] Y.Brenier: *The least action principle and the related concept of generalized flows for incompressible perfect fluids.* J. Amer. Mat. Soc., **2** (1989), 225–255.

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

[14] Y.Brenier: *The dual least action problem for an ideal, incompressible fluid.* Arch. Rational Mech. Anal., **122** (1993), 323–351.

[15] Y.Brenier: *A homogenized model for vortex sheets.* Arch. Rational Mech. Anal., **138** (1997), 319–353.

[16] Y.Brenier: *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations.* Comm. Pure Appl. Math., **52** (1999), 411–452.

[17] Y.Brenier: *Convergence of the Vlasov-Poisson system to the incompressible Euler equations.* Comm. Partial Differential Equations, **25** (2000), no. 3-4, 737–754.

[18] Y.Brenier & W.Gangbo: *L^p approximation of maps by diffeomorphisms.* Calc. Var. Partial Differential Equations, **16** (2003), no. 2, 147–164.

[19] C.Castaing & M.Valadier: *Convex Analysis and Measurable Multifunctions.* Lectures Notes in Mathematics, **580**, Spinger-Verlag, Berlin-New York, 1977.

[20] D.G.Ebin & J.Marsden: *Groups of diffeomorphisms and the motion of an ideal incompressible fluid.* Annals of Math., **2** (1970), 102–163.

[21] P.L.Lions: *Mathematical topics in fluid mechanics, Vol. I: incompressible models.* Oxford Lecture Series in Mathematics and its applications, **3**, Oxford University Press, 1996.

[22] J.Lott & C.Villani: *Ricci curvature for metric-measure spaces via optimal transport.* Annals of Math., to appear.

[23] A.I.Shnirelman: *The geometry of the group of diffeomorphisms and the dynamics of an ideal incompressible fluid.* (Russian) Mat. Sb. (N.S.), **128 (170)** (1985), no. 1, 82–109.

[24] A.I.Shnirelman: *Generalized fluid flows, their approximation and applications.* Geom. Funct. Anal., **4** (1994), no. 5, 586–620.

[25] S.K.Smirnov: *Decomposition of solenoidal vector charges into elementary solenoids and the structure of normal one-dimensional currents.* St. Petersburg Math. J., **5** (1994), 841–867.

[26] C. Villani: *Topics in optimal transportation.* Graduate Studies in Mathematics, **58**, American Mathematical Society, Providence, RI, 2003.

[27] C.Villani: *Optimal transport, old and new.* Notes from the Saint-Flour 2005 Summer School. To appear in the *Grundlehren des mathematischen Wissenschaften* Springer series. Available online at www.umpa.ens-lyon.fr/~cvillani.