FROM UNBALANCED OPTIMAL TRANSPORT TO THE CAMASSA-HOLM EQUATION

THOMAS GALLOUËT AND FRANÇOIS-XAVIER VIALARD

Abstract. The group of diffeomorphisms of a compact manifold endowed with the L^2 metric acting on the space of probability densities gives a unifying framework for the incompressible Euler equation and the theory of optimal mass transport. Recently, several authors have extended optimal transport to the space of positive Radon measures where the Wasserstein-Fisher-Rao distance is a natural extension of the classical L^2-Wasserstein distance. In this paper, we show a similar relation between this unbalanced optimal transport problem and the H^{div} right-invariant metric on the group of diffeomorphisms, which corresponds to the Camassa-Holm (CH) equation in one dimension. On the optimal transport side, we prove a polar factorization theorem on the automorphism group of half-densities. Geometrically, we present an isometric embedding of the group of diffeomorphisms endowed with this right-invariant metric in the automorphisms group of the fiber bundle of half densities endowed with an L^2 type of cone metric. This leads to a new formulation of the (generalized) CH equation as a geodesic equation on an isotropy subgroup of this automorphisms group; On S_1, solutions to the standard CH thus give particular solutions of the incompressible Euler equation on a group of homeomorphisms of $\mathbb{R}^2$ which preserve a radial density that has a singularity at 0. An other application consists in proving that smooth solutions of the Euler-Arnold equation for the H^{div} right-invariant metric are length minimizing geodesics for sufficiently short times.

1. Introduction

In his seminal article [2], Arnold showed that the incompressible Euler equation can be viewed as a geodesic flow on the group of volume preserving diffeomorphisms of a Riemannian manifold M. His formulation had an important impact in the mathematical literature and it has led to many different works. Among others, let us emphasize two different points of view which have proven to be successful.

The first one has been investigated by Ebin and Marsden in [19] where the authors have taken an intrinsic point of view on the group of diffeomorphisms as an infinite dimensional weak Riemannian manifold. Formulating the geodesic equation as an ordinary differential equation in a Hilbert manifold of Sobolev diffeomorphisms, they proved, among others, local well-posedness of the geodesic equation for smooth enough initial conditions. Since then, many fluid dynamic equations, including the Camassa-Holm equation, have been written as a geodesic flow on a group of diffeomorphisms endowed with a right-invariant metric or connection [35, 30, 48, 21, 29] and analytical properties have been derived in the spirit of [19]. Note in particular that all these works assume a strong ambient topology such as H^s for s high enough and the topology given by the Riemannian metric is generically weaker, typically L^2 in the case of incompressible Euler.

Another point of view, motivated by the variational interpretation of geodesics as minimizers of the action functional, was initiated by Brenier. He developed an extrinsic approach by considering the group of volume preserving diffeomorphisms as a Riemannian submanifold embedded in the space of maps $L^2(M, M)$ which is particularly simple when M is the Euclidean space or torus. In particular, his polar factorization theorem [5] was motivated by a numerical scheme approximating geodesics on the group of volume preserving diffeomorphisms. Optimal transport then appeared as a key tool to project a map onto this group by minimizing the L^2 distance and it can be interpreted as a non-linear extension of the pressure in the incompressible Euler equation. Since then, optimal transport has witnessed an impressive development and found many important applications inside and outside mathematics, see for instance the gigantic monograph of Villani [57]. Brenier also used optimal transport in order to define the notion of generalized geodesics for the incompressible Euler equation in [6].

In this article, we develop Brenier's point of view for a generalization in any dimension of the Camassa-Holm equation. Indeed, we present an isometric embedding of the group of diffeomorphisms endowed with the right-invariant H^{div} metric into a space of maps endowed with an L^2 metric. Moreover, the recently introduced Wasserstein-Fisher-Rao distance [13, 12], a generalization of optimal transport to measures that do not have the same total mass, plays the role of the L^2 Wasserstein distance for the incompressible Euler equation.

Before presenting our contributions in more details, we give a brief overview of the link between optimal transport and the incompressible Euler equation hereafter.

1.1. **Optimal transport and the incompressible Euler equation.** We first start from the usual static formulation of optimal transport and then present the dynamical formulation proposed by Benamou and Brenier. The link between the two formulations can be introduced via Otto's Riemannian submersion, which also provides a clear connection between incompressible Euler equation and the dynamical formulation of optimal transport. Our presentation closely follows the discussion in [32, Appendix A.5] and interesting complements can be found in [50, 30, 31]. In the rest of the section, unless otherwise mentioned, M denotes the flat torus; However, most of the results discussed hereafter are valid on Riemannian manifolds and also on much more general spaces.

Static formulation of optimal mass transport: The optimal mass transport problem as introduced by Monge in 1781 consists in finding, between two given probability measures ν_1 and ν_2, a map φ such that $\varphi_* \nu_1 = \nu_2$, i.e. the image measure of ν_1 by φ is equal to ν_2 and which minimizes a cost given by

$$\int_M c(x, \varphi(x))\, \mathrm{d}\nu_1(x)\,, \tag{1.1}$$

where c is a positive function that represents the cost of moving a particule of unit mass from location x to location y. This problem is ill-posed in the sense that solutions may not exist and the Kantorovich formulation of the problem is the correct relaxation of the Monge formulation, which can be presented as follows: On the space of probability measures on the product space $M \times M$, denoted by $\mathcal{P}(M \times M)$, find a minimizer to

$$\mathcal{I}(m) = \int_{M^2} c(x,y)\, \mathrm{d}m(x,y) \text{ such that } p^1_*(m) = \nu_1 \text{ and } p^2_*(m) = \nu_2\,, \tag{1.2}$$

where $p^1_*(m), p^2_*(m)$ denote respectively the image measure of $m \in \mathcal{P}(M \times M)$ under the projections on the first and second factors on $M \times M$. Most often in the litterature, the cost c is chosen as a power of a distance. From now on, we will only discuss the case $c(x,y) = d(x,y)^2$ where d is the distance associated with a Riemannian metric on M. In this case, the Kantorovich minimization problem defines the so-called L^2-Wasserstein distance on the space of probability measures. The Monge formulation can be expressed as a minimization problem as follows

$$W_2(\mu,\nu)^2 \stackrel{\text{def.}}{=} \inf_{\varphi \in \mathrm{Diff}(M)} \left\{ \int_M d(\varphi(x), x)^2 \; \mathrm{d}\nu_1(x) \,:\, \varphi_* \nu_1 = \nu_2 \right\}, \tag{1.3}$$

where $\mathrm{Diff}(M)$ denotes the group of smooth diffeomorphisms of M.

Dynamic formulation: In [3], Benamou and Brenier introduced a dynamical version of optimal transport which was inspired and motivated by the study of the incompressible Euler equation. Let $\rho \in C^\infty(M, \mathbb{R}_+)$ be a positive function, note that all the quantities will be implicitly time dependent. The dynamic formulation of the Wasserstein distance consists in minimizing

$$\mathcal{E}(v) = \int_0^1 \int_M \|v(t,x)\|^2 \rho(t,x) \; \mathrm{dvol}(x)\, \mathrm{d}t \;\;, \tag{1.4}$$

subject to the constraints $\dot\rho + \mathrm{div}(v\rho) = 0$ and initial condition $\rho(0) = \rho_0$ and final condition $\rho(1) = \rho_1$. The notation $\| \cdot \|$ stands for the Euclidean norm.

Equivalently, following [3], a convex reformulation using the momentum $\mathsf{m} = \rho v$ reads

$$\mathcal{E}(\mathsf{m}) = \int_0^1 \int_M \frac{\|\mathsf{m}(t,x)\|^2}{\rho(t,x)} \,\mathrm{dvol}(x)\,\mathrm{d}t\,, \tag{1.5}$$

subject to the constraints $\dot{\rho}+\mathrm{div}(\mathsf{m}) = 0$ and initial condition $\rho(0) = \rho_0$ and final condition $\rho(1) = \rho_1$. Let us underline that the functional $\mathcal{E}$ is convex in ρ, m and the continuity equation is linear in (ρ, m), therefore convex optimization methods can be applied for numerical purposes. Due to the continuity equation, the problem is feasible if and only if the initial and final densities have the same total mass using Moser's lemma [51].

Otto's Riemannian submersion: The link between the static and dynamic formulations is made clear using Otto's Riemannian submersion [52] which emphasizes the idea of a group action on the space of probability densities. Let $\mathrm{Dens}_p(M)$ be the set of probability measures that have smooth positive densities with respect to the volume measure vol. We consider such a probability density denoted by ρ_0. Otto showed that the map

$$\begin{aligned} \pi : \mathrm{Diff}(M) &\to \mathrm{Dens}_p(M) \\ \pi(\varphi) &= \varphi_*(\rho_0) \end{aligned}$$

is a formal Riemannian submersion of the metric $L^2(\rho_0)$ on $\mathrm{Diff}(M)$ to the L^2-Wasserstein metric on $\mathrm{Dens}_p(M)$. For a brief reminder on Riemannian submersions, we refer the reader to Appendix A.1. The fiber of this Riemannian submersion at point $\rho_0 \equiv 1$ is the subgroup of diffeomorphisms preserving the volume measure vol, we denote it by $\mathrm{SDiff}(M)$ and we denote its tangent space at Id by $\mathrm{SVect}(M)$, the space of divergence free vector fields. The vertical space at a diffeomorphism $\varphi \in \mathrm{Diff}(M)$ for $\rho \stackrel{\text{def.}}{=} \varphi_*\rho_0$ is

$$\mathrm{Vert}_\varphi = \{v \circ \varphi\,;\, v \in \mathrm{Vect}(M) \text{ s.t. } \mathrm{div}(\rho v) = 0\}\,. \tag{1.6}$$

In particular, consider $\varphi \in \mathrm{SDiff}(M)$, the vertical space is $\mathrm{Vert}_\varphi = \{v \circ \varphi\,;\, v \in \mathrm{SVect}(M)\}$. and the horizontal space is

$$\mathrm{Hor}_\varphi = \{\nabla p \circ \varphi\,;\, p \in C^\infty(M, \mathbb{R})\}\,. \tag{1.7}$$

Incompressible Euler equation: On the fiber $\mathrm{SDiff}(M)$, the $L^2(\mathrm{vol})$ metric is right-invariant. In Arnold's seminal work [2], it is shown that the Euler-Lagrange equation associated with this metric is the incompressible Euler equation. Arnold derived this equation as a particular case of geodesic equations on a Lie group endowed with a right-invariant metric. In its Eulerian formulation, the incompressible Euler equation is, when $M = \mathbb{T}^d$ the flat torus for the Lebesgue measure,

$$\begin{cases} \partial_t v(t,x) + v(t,x)\cdot\nabla v(t,x) = -\nabla p(t,x), \quad t > 0,\ x \in M\,, \\ \mathrm{div}(v) = 0\,, \\ v(0,x) = v_0(x)\,, \end{cases} \tag{1.8}$$

where $v_0 \in \mathrm{SVect}(M)$ is the initial condition and p is the pressure function. On a general Riemannian manifold (M, g) compact and without boundary, the formulation is similar, omitting the time and space variables, for the volume measure,

$$\begin{cases} \partial_t v + \nabla_v v = -\nabla p, \quad t > 0,\ x \in M\,, \\ \mathrm{div}(v) = 0\,, \\ v(0,x) = v_0(x)\,, \end{cases} \tag{1.9}$$

where, in this case, the symbol ∇ denotes the Levi-Civita connection associated with the Riemannian metric on M and div denotes the divergence w.r.t. the volume measure. Another fruitful point of view consists in considering the group $\mathrm{SDiff}(M)$ as isometrically embedded in the group $\mathrm{Diff}(M)$ endowed with the $L^2(\mathrm{vol})$ (non right-invariant) metric. Therefore the geodesic equations are simply geodesic equations on the Riemannian submanifold $\mathrm{SDiff}(M)$ and the geodesic equations can be written as

$$\ddot{\phi} = -\nabla p \circ \phi \,, \tag{1.10}$$

where $\phi \in \mathrm{SDiff}(M)$ and p is still a pressure function. Using this Riemannian submanifold approach, Brenier was able to prove that solutions for which the Hessian of p is bounded in L^∞ are length minimizing for short times and several of his analytical results were derived from this formulation [4, 6].

Inviscid Burgers equation: The geodesic equation on the group of diffeomorphisms for the L^2 metric written in Eulerian coordinates is the compressible Burgers equation. Its formulation on $M = \mathbb{T}^d$ is

$$\partial_t u(t,x) + u(t,x) \cdot \nabla u(t,x) = 0 \,, \tag{1.11}$$

or on a general Riemannian manifold

$$\partial_t u + \nabla_u u = 0 \,. \tag{1.12}$$

This formulation is obviously related to the incompressible Euler equation where the pressure p can be interpreted as a Lagrange multiplier associated with the incompressibility constraint, which is not present in Burgers equation. Since the map π is a Riemannian submersion, geodesics on the space of densities can be lifted horizontally to geodesics on the group. These horizontal geodesics are potential solutions of the Burgers equation, if $u_0 = \nabla q_0$, i.e. u is a potential at the initial time, then u_t stays potential for all time (until it is not well defined any longer). The corresponding equation for the potential q is the Hamilton-Jacobi equation

$$\partial_t q(t,x) + \frac{1}{2}\|\nabla q(t,x)\|^2 = 0 \,, \tag{1.13}$$

which, in this formulation, makes sense on a Riemannian manifold.

1.2. **Previous works and contributions.** Recently, several authors including the second author extended optimal transport to the case of unbalanced measures, i.e. measures that do not have the same total mass. Although several works extended optimal transport to this setting, surprisingly enough, the equivalent of the L^2-Wasserstein distance in this unbalanced setting has been introduced in 2015 simultaneously by [13, 12] motivated by imaging applications, [38, 39] motivated by gradient flows as well as [34] and by [53] for optimal transport of contact structures. In this paper, we show that, in the case of the Wasserstein-Fisher-Rao metric, the equivalent to the incompressible Euler equation is a generalization of the Camassa-Holm equation, namely the Euler-Arnold equation for the right-invariant metric H^{div} on the group of diffeomorphisms. In one dimension, geodesics for the right-invariant H^{div} metric are the solutions to the Camassa-Holm equation introduced in [11]. Since its introduction, the Camassa-Holm equation has attracted a lot of attention since it is a bi-Hamiltonian system as well as an integrable system, it exhibits peakon solutions and it is a model for waves in shallow water [16, 14, 37, 15, 8, 17, 28]. In particular, this equation is known for its well understood blow-up in finite time and is a model for wave breaking [43].

Although the title of [9], which refers to optimal transport and the Camassa-Holm equation, is seemingly close to our article, the authors introduce a metric based on optimal transport which gives Lipschitz estimates for the solutions of the Camassa-Holm equation and it is a priori completely different to our construction. Indeed, in our article, the optimal transport metric measures the discrepancy of not being in the stabilizer of the group action defined in Section 2.3 where the solutions of the Camassa-Holm equation lie.

Two complementary directions are developed in this article:

First, we study in details the associated optimal transport problem on Riemannian manifolds. In particular, we prove an equivalence between the dynamic and static formulations of the unbalanced optimal transport problem on Riemannian manifolds and we provide a detailed study of the dual problem. This enables us to formulate a generalization of the polar factorization on the automorphism group of the fiber bundle of half-densities.

Second, we rewrite the geodesic flow of the right-invariant H^{div} metric on the diffeomorphism group as a geodesic equation on a constrained submanifold of a semidirect product of group or equivalently on the automorphism group of the half-densities fibre bundle endowed with the cone metric (see Section 2.2 for its definition). This point of view has three applications: (1) We interpret solutions to the Camassa-Holm equation and one of its generalization in higher dimension as particular solutions of the incompressible Euler equation on the plane for a radial density which has a singularity at 0. This correspondence can be introduced via a sort of Madelung transform. (2) We generalize a result of Khesin et al. in [30] by computing the curvature of the group as a Riemannian submanifold. (3) Generalizing a result of Brenier to the case of Riemannian manifolds, which states that solutions of the incompressible Euler equation are length minimizing geodesic for sufficiently short times, we prove similar results for the Camassa-Holm equation.

1.3. **Plan of the paper.** In Section 2, we introduce the Wasserstein-Fisher-Rao metric which generalizes the L^2 Wasserstein metric on the space of *probability* densities to the space of densities, thus relaxing the mass constraint. Its presentation emphasizes the dynamical formulation, similar to the Benamou-Brenier formulation. This dynamical formulation naturally introduces a cone metric which is detailed in Section 2.2. Then, we present the generalization of Otto's Riemannian submersion to this unbalanced case. This generalization uses a semidirect product of group which can be interestingly interpreted as the automorphism group of the principal fibre bundle of half-densities, as explained in Section 2.3. This semidirect product of group has a natural left action on the space of densities and it gives the Riemannian submersion between an L^2 type of metric on the group and the Wasserstein-Fisher-Rao metric on the space of densities. This group action point of view is a key point to introduce, in Section 3, the Monge formulation of the Wasserstein-Fisher-Rao metric and we also prove, in Section 3.2, that the dynamic formulation is equivalent to its Kantorovich counterpart in the case of a compact Riemannian manifold. This is done by a direct generalization of the arguments in [42]. In Section 3, we also propose a generalization of Brenier's polar factorization in this context which can be interpreted as a constrained version of it on the automorphism group of the half-densities principal fibre bundle.

In Section 4, we explain the Euler-Arnold derivation of the incompressible Euler equation and other fluid dynamic equations such as the Camassa-Holm equation. We then recall in Section 4.3 the Ebin-Marsden approach to show local well-posedness of the Camassa-Holm equation.

Section 5 presents the corresponding submanifold point of view corresponding to the Camassa-Holm equation (its generalization). The submanifold is the isotropy subgroup of the left action of the semidirect product of group and the ambient metric is the L^2 type of metric. As a direct consequence, it gives a generalization of a result on the sectional curvature written in [30, Theorem A.2].

Two main applications are detailed in Section 6. In Section 6.1, we show that solutions of the Camassa-Holm equation (its generalization) can be seen as particular solutions of an incompressible Euler equation for a particular density on the cone which has a singularity at 0. In Section 6.2, we improve a result of Ebin and Marsden by applying Brenier's approach to show that every smooth geodesics are length minimizing on a sufficiently short time interval under mild conditions.

1.4. **Notations.** Hereafter is a non exhaustive list of notations used throughout the paper.

- (M, g) is a smooth orientable Riemannian manifold which is assumed compact and without boundary. Its volume form is denoted by vol, TM and T^*M denote respectively the tangent and the cotangent bundle.
- The distance on (M, g) is sometimes denoted by d_M when a confusion might occur.
- For $x \in M$, the squared norm of a vector $v \in T_xM$ will be denoted by $\|v\|^2$ or $g(x)(v, v)$.
- For $x \in M$, we denote by $\exp_x^M : T_xM \to M$, the exponential map, the superscript being a reminder of the underlying manifold.
- $\mathcal{C}(M)$ is the Riemannian cone over (M, g) and is introduced in Definition 2.
- The operator div is the divergence w.r.t. the volume form on (M, g).
- The Lie bracket between two vector fields X, Y on M is denoted by $[X, Y]$.

- If $f \in C^1(M,\mathbb{R})$, then ∇f is the gradient of f w.r.t. the metric g. Sometimes, we use the notation ∇_x to make clear which variable we consider.
- The group of invertible linear maps on $\mathbb{R}^d$ is denoted by $\mathrm{GL}_d(\mathbb{R})$.
- For a quantity $f(t,x)$ that depends on time and space variable, we denote by $\dot{f}$ its time derivative.
- On $\mathbb{R}$ and $\mathbb{C}$, $|\cdot|$ denotes respectively the absolute value and the module.
- $M = S_n(r)$ the Euclidean sphere of radius r in $\mathbb{R}^{n+1}$.
- The Lebesgue measure is denoted by Leb.
- Sometimes, we use the notation $a \overset{\text{def.}}{=} b$ to define a as b.

2. A Geometric Point of View on Unbalanced Optimal Transport

2.1. The Wasserstein-Fisher-Rao metric, its dynamical formulation. The continuity equation enforces the mass conservation property in the Benamou-Brenier formulation (1.4). This constraint can be relaxed by introducing a source term $\mu \in C^\infty(M,\mathbb{R})$,

$$\dot{\rho} = -\operatorname{div}(\rho v) + \mu\,. \tag{2.1}$$

For a given variation of the density $\dot{\rho}$, there exist a priori many couples (v,μ) that reproduce this variation. Following [55], it can be determined via the minimization of the norm of (v,μ), for a given choice of norm. The penalization of μ was chosen in [41] as the L^2 norm but a natural choice is rather the Fisher-Rao metric

$$\mathrm{FR}^2(\mu) = \int_M \frac{\mu(t,x)^2}{\rho(t,x)}\,\mathrm{dvol}(x)\,,$$

because it is homogeneous. In other words, this is the L^2 norm of the growth rate w.r.t. the density ρ since it can be written as $\int_M \alpha(t,x)^2\rho(t,x)\,\mathrm{dvol}(x)$ where α is the growth rate $\alpha(t,x) \overset{\text{def.}}{=} \frac{\mu(t,x)}{\rho(t,x)}$. Note in particular that this action is 1-homogeneous with respect to the couple (μ,ρ). This point is important for convex analysis properties and especially, in order to define the action functional on singular measures via the same formula. Obviously, there are many other choices of norms that satisfies this homogeneity property but this particular one can be related to the Camassa-Holm equation.

Thus, the Wasserstein-Fisher-Rao functional also known as Hellinger-Kantorovich [38], or Kantorovich-Fisher-Rao [26], is simply given by the infimal convolution between the Wasserstein and the Fisher-Rao metric tensor.

Definition 1 (WF metric)**.** Let (M,g) be a smooth Riemannian manifold compact and without boundary, $a,b \in \mathbb{R}_+^*$ be two positive real numbers and $\rho_0,\rho_1 \in \mathcal{M}_+(M)$ be two nonnegative Radon measures. The Wasserstein-Fisher-Rao metric is defined by

$$\mathrm{WF}^2(\rho_0,\rho_1) = \inf_{\rho,\mathsf{m},\mu} \mathcal{J}(\rho,\mathsf{m},\mu)\,, \tag{2.2}$$

where

$$\mathcal{J}(\rho,\mathsf{m},\mu) = a^2\int_0^1\int_M \frac{g^{-1}(x)(\tilde{\mathsf{m}}(t,x),\tilde{\mathsf{m}}(t,x))}{\tilde{\rho}(t,x)}\,\mathrm{d}\nu(t,x)\; + b^2\int_0^1\int_M \frac{\tilde{\mu}(t,x)^2}{\tilde{\rho}(t,x)}\,\mathrm{d}\nu(t,x) \tag{2.3}$$

over the set (ρ,m,μ) satisfying $\rho \in \mathcal{M}([0,1]\times M)$, $\mathsf{m} \in (\Gamma_M^0([0,1]\times M, TM))^*$ which denotes the dual of time dependent continuous vector fields on M (time dependent sections of the tangent bundle), $\mu \in \mathcal{M}([0,1]\times M)$ subject to the constraint

$$\int_0^1\int_M \partial_t f\,\mathrm{d}\rho + \int_0^1\int_M \mathsf{m}(\nabla_x f) - f\mu\,\mathrm{d}\nu = \int_M f(1,\cdot)\,\mathrm{d}\rho_1 - \int_M f(0,\cdot)\,\mathrm{d}\rho_0 \tag{2.4}$$

satisfied for every test function $f \in C^1([0,1]\times M,\mathbb{R})$. Moreover, ν is chosen such that ρ,m,μ are absolutely continuous with respect to ν and $\tilde{\rho},\tilde{\mathsf{m}},\tilde{\mu}$ denote their Radon-Nikodym derivative with respect to ν.

Remark 1. *Note that, in the previous definition, the divergence operator* $\operatorname{div}(\cdot)$ *is defined by duality on the space of* C^1 *functions. In addition, since the functions in the integrand of formula (2.2) are one homogeneous with respect to the triple of arguments* $(\tilde{\rho}, \tilde{\mathsf{m}}, \tilde{\mu})$*, the functional does not depend on the choice of* ν *which dominates the measures. Last, the Radon-Nikodym theorem applied to the measure* m *gives* $\mathsf{m} = \tilde{\mathsf{m}}\nu$ *where* $\tilde{\mathsf{m}}$ *is a measurable section of* T^*M*.*

The following property is immediate to check.

Proposition 1. *The* WF^2 *functional is convex and positively one homogeneous on the space of Radon measures. Moreover,* WF *defines a distance on the space of nonnegative Radon measures which is continuous w.r.t. to the weak-* topology.*

This dynamical formulation enjoys most of the analytical properties of the initial Benamou-Brenier formulation (1.4) and especially convexity. An important consequence is the existence of optimal paths in the space of time-dependent measures [13] by application of the Fenchel-Rockafellar duality theorem as stated in the next proposition. We omit the proof here since it is similar to the one given in [12] in the Euclidean case and it is also proven in more general spaces in [38].

Proposition 2 (Hamilton-Jacobi)**.** *There exists a minimum to the minimization problem (2.2) and it holds*

$$\mathrm{WF}^2(\rho_0, \rho_1) = \sup_{q \in C} \int_M q(1, \cdot)\,\mathrm{d}\rho_1 - \int_M q(0, \cdot)\,\mathrm{d}\rho_0 \tag{2.5}$$

where C *is the set of functions* $q \in C^1([0,1] \times M, \mathbb{R})$ *such that*

$$\partial_t q(t,x) + \frac{1}{2a^2}\|\nabla q(t,x)\|^2 + \frac{1}{2b^2} q(t,x)^2 \leq 0\,. \tag{2.6}$$

Not only analytical properties of standard optimal transport are conserved but also some interesting geometrical properties such as the Riemannian submersion highlighted by Otto, as explained in the introduction. More precisely, the group of diffeomorphisms can be replaced by a semi-direct product of group between $\mathrm{Diff}(M)$ and the space $C^\infty(M, \mathbb{R}_+^*)$ which is a group under pointwise multiplication. In addition, this group acts on the space of densities $\mathrm{Dens}(M)$ and this action gives a Riemannian submersion between the group endowed with an L^2 type of metric, namely $L^2(M, \mathcal{C}(M))$ and the space of densities endowed with the Wasserstein-Fisher-Rao metric. The notation $\mathcal{C}(M)$ is the cone over M defined in the next section 2.2, it is the manifold $M \times \mathbb{R}_+^*$ endowed with the Riemannian metric given in Definition 2. Moreover, this semidirect product of groups is naturally identified as the automorphism group of the fibre bundle of half densities in section 2.3.

2.2. **A cone metric.** To motivate the introduction of the cone metric, let us first discuss informally what happens for a particle of mass $m(t)$ at a spatial position $x(t)$ in a Riemannian manifold (M, g) under the generalized continuity constraint (2.1). Let us assume the following structure for the measure $m(t)\delta_{x(t)}$ where $m(t) \in \mathbb{R}_+^*$ is the mass of the Dirac measure and $x(t) \in M$ its location, the system reads

$$\begin{cases} \dot{x}(t) = v(x(t)) \\ \dot{m}(t) = \alpha(x(t)) m(t) \end{cases} \tag{2.7}$$

where $\alpha = \frac{\mu}{\rho}$ is the growth rate. The action associated with formula (2.2) reads $\int_0^1 a^2 |v(x(t))|^2 m(t) + b^2 \frac{\dot{m}(t)^2}{m(t)}\,\mathrm{d}t$. Thus, considering the particle as a point in $M \times \mathbb{R}_+^*$, the Riemannian metric seen by the particle is $a^2 m g + b^2 \frac{\mathrm{d}m^2}{m}$. Therefore, it will be of importance to study this Riemannian metric $M \times \mathbb{R}_+^*$. Actually, this space is isometric to the standard Riemannian cone defined below.

Definition 2 (Cone)**.** Let (M, g) be a Riemannian manifold. The cone over M denoted by $\mathcal{C}(M)$ is the quotient space $(M \times \mathbb{R}_+) \,/\, (M \times \{0\})$. The cone point $M \times \{0\}$ is denoted by $\mathcal{S}$. The cone will be endowed with the metric $g_{\mathcal{C}(M)} \stackrel{\text{def.}}{=} r^2 g + \mathrm{d}r^2$ defined on $M \times \mathbb{R}_+^*$ and r is the variable in $\mathbb{R}_+^*$.

The isometry is given by the square root change of variable on the mass, as stated in the following proposition.

Proposition 3. *The space* $(M \times \mathbb{R}_+^*, mg + \frac{1}{4m}\,\mathrm{d}m^2)$ *is isometric to* $(\mathcal{C}(M), g_{\mathcal{C}(M)})$ *by the change of variable* $r = \sqrt{m}$*. Therefore, the distance on* $(M \times \mathbb{R}_+^*, a^2 mg + \frac{b^2}{m}\,\mathrm{d}m^2)$ *is given by*

$$d((x_1, m_1), (x_2, m_2))^2 = 4b^2 \left(m_2 + m_1 - 2\sqrt{m_1 m_2} \cos\left(\frac{a}{2b} d_M(x_1, x_2) \wedge \pi\right)\right) . \tag{2.8}$$

Moreover, if c *is a geodesic for the metric* $\frac{a^2}{4b^2} g$*, an isometry* $S : \mathbb{C} \setminus \mathbb{R}_- \to M \times \mathbb{R}_+^*$ *is defined by* $S(\sqrt{m} e^{i\theta}) = (c(\theta), 2bm)$*.*

In physical terms, it implies that mass can "appear" and "disappear" at finite cost. In other words, the Riemannian cone is not complete but adding the cone point, which represents $M \times \{0\}$, to $M \times \mathbb{R}_+^*$ turns it into a complete metric space when M is complete. Importantly, the distance associated with the cone metric (2.8) is 1-homogeneous in (m_1, m_2). In the rest of the paper, unless explicitly mentioned, we consider the case $a = 1$ and $b = 1/2$. We now collect known facts about Riemannian cones.

Proposition 4. *On the cone* $\mathcal{C}(M)$*, we denote by* e *the vector field defined by* $\frac{\partial}{\partial m}$*. The Levi-Civita connection on* (M, g) *will be denoted by* ∇^g*. For a given vector field* X *on* M*, define its lift as a vector field on* $M \times \mathbb{R}_+^*$ *by* $\hat{X}(x, r) = (X(x), 0)$*. The Levi-Civita connection on* $\mathcal{C}(M)$ *denoted by* ∇ *is given by*

$$\nabla_{\hat{X}} \hat{Y} = \widehat{\nabla^g_X Y} - r g(X, Y) e\,, \;\; \nabla_e e = 0 \text{ and } \nabla_e \hat{X} = \nabla_{\hat{X}} e = \frac{1}{r} \hat{X}\,.$$

The curvature tensor R *on the cone satisfies the following properties,*

$$R(\hat{X}, e) = 0 \text{ and } R(\hat{X}, \hat{Y})\hat{Z} = (R_g(X, Y)Z - g(Y, Z)X + g(X, Z)Y, 0) \tag{2.9}$$

where R_g *denotes the curvature tensor of* (M, g)*. Let* X, Y *be two orthornormal vector fields on* M*,*

$$K(\hat{X}, \hat{Y}) = K_g(X, Y) - 1 \tag{2.10}$$

where K *and* K_g *denote respectively the sectional curvatures of* $\mathcal{C}(M)$ *and* M*.*

Proof. Direct computations, see [23]. □

Let us give simple comments on Riemannian cones: Usual cones, embedded in $\mathbb{R}^3$ are cones over S_1 of length less than 2π. Although Riemannian cones over a segment in $\mathbb{R}$ are locally flat, the curvature still concentrates at the cone point. The cone over the sphere is isometric to the Euclidean space (minus the origin) and the cone over the Euclidean space has nonpositive curvature. In particular, the cone over S_1 is isometric to $\mathbb{R}^2 \setminus \{0\}$. We refer to [10] for more informations on cones from the point of view of metric geometry.

We need the explicit formulas for the geodesic equations on the cone.

Corollary 5. *The geodesic equations on the cone* $\mathcal{C}(M)$ *are given by*

$$\frac{D}{Dt}^g \dot{x} + 2\frac{\dot{r}}{r}\dot{x} = 0 \tag{2.11a}$$

$$\ddot{r} - r g(\dot{x}, \dot{x}) = 0\,, \tag{2.11b}$$

where $\frac{D}{Dt}^g$ *is the covariant derivative associated with* (M, g)*.*
Alternatively, the geodesic equations on $(M \times \mathbb{R}_+^*, a^2 mg + \frac{b^2}{m}\,\mathrm{d}m^2)$ *can be written w.r.t. the initial "mass" coordinate as follows*

$$\frac{D}{Dt}^g \dot{x} + \frac{\dot{m}}{m}\dot{x} = 0 \tag{2.12a}$$

$$\ddot{m} - \frac{\dot{m}^2}{2m} - \frac{a^2}{2b^2} g(\dot{x}, \dot{x}) m = 0\,. \tag{2.12b}$$

2.3. **The automorphism group of the bundle of half-densities.** The cone can be seen as a trivial principal fibre bundle since $\mathcal{C}(M)$ is the direct product of M with the group $\mathbb{R}_+^*$. Let us denote $p_M : \mathcal{C}(M) \mapsto M$ the projection on the first factor. The group $\mathbb{R}_+^*$ induces a group action on $\mathcal{C}(M)$ defined by $\lambda \cdot (x, \lambda') \stackrel{\text{def.}}{=} (x, \lambda\lambda')$, for all $x \in M$ and $\lambda, \lambda' \in \mathbb{R}_+^*$. We now identify the trivial fibre bundle of half densities with the cone.

Definition 3. Let M be a smooth manifold without boundary and (U_α, u_α) be a smooth atlas. The bundle of s-densities is the line bundle given by the following cocycle

$$\Psi_{\alpha\beta} : U_\alpha \cap U_\beta \mapsto \mathrm{GL}_1(\mathbb{R}) = \mathbb{R}^*$$
$$\Psi_{\alpha\beta}(x) = |\det(\,\mathrm{d}(u_\beta \circ u_\alpha^{-1})(u_\alpha(x))|^s = \frac{1}{|\det(d(u_\alpha \circ u_\beta^{-1}))(u_\beta(x))|^s}\,.$$

We denote by $\mathrm{Dens}_s(M)$ the set of sections of this bundle and we use $\mathrm{Dens}(M)$ instead of $\mathrm{Dens}_1(M)$, the space of densities. This definition shows that this fibre bundle is also a principal fibre bundle over $\mathbb{R}_+^*$ and it will be the point of view adopted in the rest of the paper.

On any smooth manifold M, the fibre bundle of s-densities is a trivial principal bundle over $\mathbb{R}_+^*$ since there exists a smooth positive density on M. Note that this trivialization depends on the choice of this reference positive density. If one chooses such a positive density, then the 1/2-density bundle can be identified to the cone $\mathcal{C}(M)$. *Let us fix the reference volume form to be the volume measure* vol. *By this choice, we identify* $\mathrm{Dens}_{1/2}(M)$ *with the set of sections of the cone* $\mathcal{C}(M)$ *in the rest of the paper.* Thus every element of $\mathrm{Dens}_{1/2}(M)$ is a section of the cone $\mathcal{C}(M)$. We are now interested in transformations that preserve the group structure. Namely, one can define

$$\mathrm{Aut}(\mathcal{C}(M)) = \left\{\Phi \in \mathrm{Diff}(\mathcal{C}(M))\,;\, \Phi(x, r) = r \cdot \Phi(x, 1) \text{ for all } r \in \mathbb{R}_+^*\right\}\,, \tag{2.13}$$

which is the instantiation, in this particular case, of the definition of the automorphisms group of a principal fibre bundle. In other words, this is the subgroup of diffeomorphisms of the cone that preserve the group action on the fibers. In particular, $\mathrm{Aut}(\mathcal{C}(M))$ is a subgroup of $\mathrm{Diff}(\mathcal{C}(M))$. Of particular interest is the subgroup of $\mathrm{Aut}(\mathcal{C}(M))$ which is defined as

$$\mathrm{Gau}(\mathcal{C}(M)) = \{\Phi \in \mathrm{Aut}(\mathcal{C}(M))\,;\, p_M \circ \Phi = \mathrm{id}_M\}\,. \tag{2.14}$$

The set $\mathrm{Gau}(\mathcal{C}(M))$ is called the gauge group and it is a normal subgroup of $\mathrm{Aut}(\mathcal{C}(M))$. We now consider the injection $s : \mathrm{Diff}(M) \hookrightarrow \mathrm{Aut}(\mathcal{C}(M))$ defined by $s(\varphi) = (\varphi, \mathrm{id}_{\mathbb{R}_+^*})$. This is the standard situation of a semidirect product of groups between $i(\mathrm{Diff}(M))$ and $\mathrm{Gau}(\mathcal{C}(M))$ since the following sequence is exact

$$\mathrm{Gau}(\mathcal{C}(M)) \hookrightarrow \mathrm{Aut}(\mathcal{C}(M)) \to \mathrm{Diff}(M)\,, \tag{2.15}$$

where s defined above provides an associated section of the short exact sequence. Note that we could also have chosen the natural section associated to the natural bundle of half-densities. As is well-known for a trivial principal bundle, $\mathrm{Aut}(\mathcal{C}(M))$ is therefore equal to the semidirect product of group:

$$\mathrm{Aut}(\mathcal{C}(M)) = \mathrm{Diff}(M) \ltimes_\Psi \mathrm{Gau}(\mathcal{C}(M))\,, \tag{2.16}$$

where $\Psi : \mathrm{Diff}(M) \mapsto \mathrm{Aut}(\mathrm{Gau}(\mathcal{C}(M)))$ is defined by $\Psi(\varphi)(\lambda) = \varphi^{-1}\lambda\varphi$ being the associated inner automorphism of the group $\mathrm{Gau}(\mathcal{C}(M))$, where the composition is understood as composition of diffeomorphisms of $\mathcal{C}(M)$. Being a trivial principal fibre bundle, the gauge group can be identified with the space of positive functions on M. Let us denote $\Lambda_{1/2}(M) \stackrel{\text{def.}}{=} C^\infty(M, \mathbb{R}_+^*)$ which is a group under pointwise multiplication. The subscript 1/2 is a reminder of the fact that $\Lambda_{1/2}(M)$ is the gauge group of $\mathcal{C}(M)$, the bundle of 1/2-densities. Note that we do not use the standard left action but, instead, a right action for the inner automorphisms as presented in [33, Section 5.3], which fits better to our notations, although these two choices are equivalent. The identification of $\Lambda_{1/2}$ with the gauge group $\mathrm{Gau}(\mathcal{C}(M))$ is simply $\lambda \mapsto (\mathrm{id}_M, \lambda)$ where $(\mathrm{id}_M, \lambda) : (x, m) \mapsto (x, \lambda(x)m)$. The group composition law is given by

$$(\varphi_1, \lambda_1) \cdot (\varphi_2, \lambda_2) = (\varphi_1 \circ \varphi_2, (\lambda_1 \circ \varphi_2)\lambda_2) \tag{2.17}$$

and the inverse is

$$(\varphi, \lambda)^{-1} = (\varphi^{-1}, \lambda^{-1} \circ \varphi^{-1}) \,. \tag{2.18}$$

By construction, the group $\mathrm{Aut}(\mathcal{C}(M))$ has a natural left action on the space $\mathrm{Dens}_{1/2}(M)$ as well as on $\mathrm{Dens}(M)$. The action on $\mathrm{Dens}(M)$ is explicitly given by the map π defined by

$$\begin{aligned} &\pi : \big(\mathrm{Diff}(M) \ltimes_\Psi \Lambda_{1/2}(M)\big) \times \mathrm{Dens}(M) \mapsto \mathrm{Dens}(M) \\ &\pi\left((\varphi, \lambda), \rho\right) \overset{\text{def.}}{=} \varphi_*(\lambda^2 \rho) \,. \end{aligned} \tag{2.19}$$

We will also use sometimes the notation $\pi\left[(\varphi, \lambda), \rho\right]$. For particular choices of metrics, this left action is a Riemannian submersion as detailed below. Note that we will use both automorphism group and semidirect product notations equally, depending on the context.

2.4. **A Riemannian submersion between the automorphism group and the space of densities.** The semidirect product of group $\mathrm{Diff}(M) \ltimes_\Psi \Lambda_{1/2}(M)$ will be endowed with the metric $L^2(M, \mathcal{C}(M))$ with respect to the reference measure on M. This is probably the simplest type of (weak) Riemannian metrics on spaces of mappings and it has been studied in details in [22] where, in particular, the curvature is computed. Note in particular that this metric is *not* the right-invariant metric L^2 on the semidirect product of groups as in [29] or on the automorphism group which would lead to an EPDiff equation on a principal fibre bundle as developed in [27].

Proposition 6. *The geodesic equations on* $\mathrm{Aut}(\mathcal{C}(M))$ *endowed with the metric* $L^2(M, \mathcal{C}(M))$ *with respect to the reference measure on* ν *are given by the geodesic equations on the cone* (2.11)*, that is* $\frac{D}{Dt}(\dot\varphi, \dot\lambda) = 0$*, or more explicitely*

$$\frac{D}{Dt}^{g} \dot\varphi + 2\frac{\dot\lambda}{\lambda}\dot\varphi = 0 \tag{2.20a}$$

$$\ddot\lambda - \lambda g(\dot\varphi, \dot\varphi) = 0 \,. \tag{2.20b}$$

Remark 2. *Note that the first equation* (2.20a) *is* 0*-homogeneous with respect to* λ *and the second equation* (2.20b) *is one homogeneous with respect to* λ*. Therefore, the automorphism group* $\mathrm{Aut}(\mathcal{C}(M))$ *is totally geodesic in* $\mathrm{Diff}(\mathcal{C}(M))$ *for the* $L^2(\mathcal{C}(M), \mathcal{C}(M))$ *metric. This is a consequence of the fact that* $\mathbb{R}_+^*$*-multiplication acts as an affine isometry on* $\mathcal{C}(M)$*.*

Let us first recall some useful notions. From the point of view of fluid dynamics, the next definition corresponds to the change of variable between Lagrangian and Eulerian formulations.

Definition 4 (Right-trivialization)**.** Let H be a group and a smooth manifold at the same time, possibly of infinite dimensions, the right-trivialization of TH is the bundle isomorphism $\tau : TH \mapsto H \times T_{\mathrm{Id}}H$ defined by $\tau(h, X_h) \overset{\text{def.}}{=} (h, dR_{h^{-1}} X_h)$, where X_h is a tangent vector at point h and $\mathcal{R}_{h^{-1}} : H \to H$ is the right multiplication by h^{-1}, namely, $R_{h^{-1}}(f) = f h^{-1}$ for all $f \in H$.

In fluid dynamics, the right-trivialized tangent vector $dR_{h^{-1}} X_h$ corresponds to the spatial or Eulerian velocity and X_h is the Lagrangian velocity. Importantly, this right-trivialization map is continuous but not differentiable with respect to the variable h. Indeed, right-multiplication R_h is smooth, yet left multiplication is continuous and usually not differentiable, due to a loss of smoothness.

Example 7. *For the semi-direct product of groups defined above, we have*

$$\tau((\varphi, \lambda), (X_\varphi, X_\lambda)) = ((\varphi, \lambda), (X_\varphi \circ \varphi^{-1}, (X_\lambda \lambda^{-1}) \circ \varphi^{-1})) \,. \tag{2.21}$$

We will denote by (v, α) *an element of the tangent space of* $T_{(\mathrm{Id},1)} \mathrm{Diff}(M) \ltimes_\Psi \Lambda_{1/2}(M)$*.*

As an immediate consequence of Proposition 6, we write the geodesic equation in Eulerian coordinates.

Corollary 8 (Geodesic equations in Eulerian coordinates)**.** *After right-trivialization, that is under the change of variable* $v \stackrel{\text{def.}}{=} \dot{\varphi} \circ \varphi^{-1}$ *and* $\alpha \stackrel{\text{def.}}{=} \frac{\dot{\lambda}}{\lambda} \circ \varphi^{-1}$*, the geodesic equations read*

$$\begin{cases} \dot{v} + \nabla_v v + \alpha v = 0 \\ \dot{\alpha} + \langle \nabla \alpha, v \rangle + \alpha^2 - g(v,v) = 0 \,. \end{cases} \tag{2.22}$$

Recall now the infinitesimal action associated with a group action.

Definition 5 (Infinitesimal action)**.** For a smooth left action of H a Lie group on a manifold M and $q \in M$, the infinitesimal action is the map $T_{\mathrm{Id}} H \times M \mapsto TM$ defined by

$$\xi \cdot q \stackrel{\text{def.}}{=} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\exp(\xi t) \cdot q) \in T_q M \tag{2.23}$$

where $\cdot$ denotes the left action of H on M and $\exp(\xi t)$ is the Lie exponential, that is the solution to $\dot{h} = dR_h(\xi)$ and $h(0) = \mathrm{Id}$.

Example 9. *For* $\mathrm{Diff}(M) \ltimes_\Psi \Lambda_{1/2}(M)$ *acting on* $\mathrm{Dens}(M)$*, the previous definition gives* $(v,\alpha) \cdot \rho = -\operatorname{div}(v\rho) + 2\alpha\rho$*. Indeed, one has*

$$(\varphi(t), \lambda(t)) \cdot \rho = \mathrm{Jac}(\varphi(t)^{-1})(\lambda^2(t)\rho) \circ \varphi^{-1}(t) \,.$$

First recall that $\partial_t \varphi(t) = v \circ \varphi(t)$ *and* $\partial_t \lambda = \alpha\lambda(t)$*. Once evaluated at time* $t = 0$ *where* $\varphi(0) = \mathrm{Id}$ *and* $\lambda(0) = 1$*, the differentiation with respect to* φ *gives* $-\operatorname{div}(v\rho)$ *and the second term* $2\alpha\rho$ *is given by the differentiation with respect to* λ*.*

We now recall the result of [46, Claim of Section 29.21] in a finite dimensional setting. This result presents a standard construction to obtain Riemannian submersions from a transitive group action.

Proposition 10. *Consider a smooth left action of Lie group* H *on a manifold* M *which is transitive and such that for every* $\rho \in M$*, the infinitesimal action* $\xi \mapsto \xi \cdot \rho$ *is a surjective map. Let* $\rho_0 \in M$ *and a Riemannian metric* G *on* H *that can be written as:*

$$G(h)(X_h, X_h) = g(h \cdot \rho_0)(dR_{h^{-1}} X_h, dR_{h^{-1}} X_h) \tag{2.24}$$

for $g(h \cdot \rho_0)$ *an inner product on* $T_{\mathrm{Id}} H$*. Let* $X_\rho \in T_\rho M$ *be a tangent vector at point* $h \cdot \rho_0 = \rho \in M$*, we define the Riemannian metric* $\overline{g}$ *on* M *by*

$$\overline{g}(\rho)(X_\rho, X_\rho) \stackrel{\text{def.}}{=} \min_{\xi \in T_{\mathrm{Id}} H} g(\rho)(\xi,\xi) \text{ under the constraint } X_\rho = \xi \cdot \rho \,. \tag{2.25}$$

where $\xi = X_h \cdot h^{-1}$*.*

Then, the map $\pi_0 : H \to M$ *defined by* $\pi_0(h) = h \cdot \rho_0$ *is a Riemannian submersion of the metric* G *on* H *to the metric* $\overline{g}$ *on* M*.*

The formal application of this construction in our infinite dimensional situation leads to the result, stated in [12]:

Proposition 11. *Let* $\rho_0 \in \mathrm{Dens}(M)$ *and define the map*

$$\begin{aligned} &\pi_0 : \mathrm{Aut}(\mathcal{C}(M)) \to \mathrm{Dens}(M) \\ &\pi_0(\varphi, \lambda) = \varphi_*(\lambda^2 \rho_0) \,. \end{aligned}$$

Then, the map π_0 *is a Riemannian submersion of the metric* $L^2(M, \mathcal{C}(M))$ *on the group* $\mathrm{Aut}(\mathcal{C}(M))$ *to the Wasserstein-Fisher-Rao on the space of densities* $\mathrm{Dens}(M)$*.*

Note also that the fibers of the submersion are right-cosets of the subgroup H_0 in H. The proof of the previous proposition is in fact given by the change of variables associated with right-trivialization.

Let ρ_0 be a reference density, the application of Proposition 10 gives

$$
\begin{aligned}
G(\varphi,\lambda)((X_\varphi,X_\lambda),(X_\varphi,X_\lambda)) &= \int_M g(v,v)\rho\,\mathrm{d}x + \int_M \alpha^2\rho\,\mathrm{d}x \\
&= \int_M g(X_\varphi\circ\varphi^{-1},X_\varphi\circ\varphi^{-1})\varphi_*(\lambda^2\rho_0)\mathrm{d}x + \int_M (X_\lambda\lambda^{-1})^2\circ\varphi^{-1}\varphi_*(\lambda^2\rho_0)\mathrm{d}x \\
&= \int_M g(X_\varphi,X_\varphi)\lambda^2\rho_0\,\mathrm{d}x + \int_M X_\lambda^2\,\rho_0\,\mathrm{d}x\,.
\end{aligned}
$$

Therefore, the metric G is the $L^2(M,\mathcal{C}(M))$ metric with respect to the density ρ_0. This metric is a weak Riemannian metric in the sense of [19]. This is indeed a smooth Riemannian metric when restricted to $\mathrm{Diff}^s(M)\ltimes_\Psi\Lambda^s(M)$ the space of Sobolev maps of order s such that $s>d/2$ essentially because these Sobolev spaces are Hilbert algebras. The same result holds for the Wasserstein-Fisher-Rao metric, as shown in [12]. Moreover, in this particular situation, the horizontal lift (2.25) is well defined.

Proposition 12 (Horizontal lift)**.** *Let $\rho\in\mathrm{Dens}^s(\Omega)$ be a smooth density and $X_\rho\in H^s(\Omega,\mathbb{R})$ be a tangent vector at the density ρ. The horizontal lift at $(\mathrm{Id},1)$ of X_ρ is given by $(\frac{1}{2}\nabla\Phi,\Phi)$ where Φ is the solution to the elliptic partial differential equation:*

$$
-\,\mathrm{div}(\rho\nabla\Phi)+2\Phi\rho = X_\rho\,. \tag{2.26}
$$

By elliptic regularity, the unique solution Φ belongs to $H^{s+2}(M)$.

Proposition 13. *The* WF *metric is a weak Riemannian metric on* $\mathrm{Dens}^s(M)$.

The proof is written in [12] but let us explain it briefly. Denote by $L(\rho)^{-1}$ the inverse of the elliptic operator defined in Formula (2.26). The WF Riemannian metric tensor is then given by $\mathrm{WF}(\rho)(X,X)=\int_M L(\rho)^{-1}(X)X\,\mathrm{d}x$. Therefore the smoothness of $\mathrm{WF}(\rho)(X,X)$ reduces to the smoothness of $L(\rho)^{-1}$ which again reduces to that of $L(\rho)$ with respect to ρ. However, this metric does not admit a Levi-Civita connection in the sense of [45, Section 2.4], which is due to the loss of derivative of the map π. Indeed, since composition on $\mathrm{Diff}^{s+1}(M)$ is continuous and $H^s(M,\mathbb{R})$ is a Hilbert algebra, we have

Proposition 14. *Let $s>d/2+1$ and $k\in\mathbb{N}$. The following map is C^k*

$$
\begin{aligned}
&\pi_0:\mathrm{Diff}^{s+k+1}(M)\ltimes_\Psi\Lambda^{s+k}_{1/2}(M)\mapsto\mathrm{Dens}^s(M)\\
&\pi_0(\varphi,\lambda)=\varphi_*(\lambda^2\rho_0)\,.
\end{aligned}
$$

Unfortunately, the map π_0 for $k=0$ is only continuous due to this loss of derivatives and therefore it is not a proper Riemannian submersion in this context. To make it a proper Riemmanian submersion, one could work with Fréchet spaces. Yet, the horizontal lift can be defined on C^1 curves.

Proposition 15. *Let $c:[0,1]\to\mathrm{Dens}^s(M)$ a C^1 curve then any horizontal lift $\tilde{c}:[0,1]\to\mathrm{Diff}^{s+1}(M)\ltimes_\Psi\Lambda^s_{1/2}(M)$ is C^1.*

Proof. The horizontal lift is given by the curve on the group $\mathrm{Diff}^{s+1}(M)\ltimes_\Psi\Lambda^s_{1/2}(M)$ defined by

$$
\begin{cases}
(\varphi,\lambda)=(\varphi_0,\lambda_0)\,,\\
(\dot\varphi,\dot\lambda)=L(c(t))^{-1}(\dot c)\circ(\varphi,\lambda)\,.
\end{cases} \tag{2.27}
$$

Since the operator $L(\rho)^{-1}(\dot c)$ is smooth with respect to c, the result follows since composition is continuous on $\mathrm{Diff}^{s+1}(M)\ltimes_\Psi\Lambda^s_{1/2}(M)$. □

Let us now detail the horizontal spaces and vertical spaces at $(\varphi,\lambda)\in\mathrm{Diff}(M)\ltimes_\Psi\Lambda_{1/2}(M)$ such that $\varphi_*(\lambda^2\rho_0)=\rho$,

$$
\mathrm{Vert}_{(\varphi,\lambda)}=\{(v,\alpha)\circ(\varphi,\lambda)\,;\,(v,\alpha)\in\mathrm{Vect}(M)\times C^\infty(M,\mathbb{R})\text{ s.t. }\ \mathrm{div}(\rho v)=2\alpha\rho\}\,, \tag{2.28}
$$

and the horizontal space is

$$\mathrm{Hor}_{(\varphi,\lambda)} = \left\{ \left(\frac{1}{2}\nabla p, p \right) \circ (\varphi, \lambda)\,;\, p \in C^\infty(M,\mathbb{R}) \right\} . \tag{2.29}$$

A direct application of this Riemannian submersion viewpoint is the formal computation of the sectional curvature of the Wasserstein-Fisher-Rao in this smooth setting by applying O'Neill's formula recalled in appendix, see [12]. To recall it hereafter, we need the Lie bracket of right-invariant vector fields on $\mathrm{Diff}(M) \ltimes_\Psi \Lambda_{1/2}(M)$.

Proposition 16. *Let (v_1, α_1) and (v_2, α_2) be two tangent vectors at identity in $\mathrm{Diff}(M) \ltimes_\Psi \Lambda_{1/2}(M)$. Then,*

$$[(v_1, \alpha_1), (v_2, \alpha_2)] = ([v_1, v_2], \nabla \alpha_1 \cdot v_2 - \nabla \alpha_2 \cdot v_1) \ , \tag{2.30}$$

where $[v_1, v_2]$ denotes the Lie bracket of vector fields.

Note that the application of this formula to horizontal vector fields gives $[(\frac{1}{2}\nabla\Phi_1, \Phi_1), (\frac{1}{2}\nabla\Phi_2, \Phi_2)] = (\frac{1}{4}[\nabla\Phi_1, \nabla\Phi_2], 0)$.

Proposition 17. *Let ρ be a smooth positive density on M and X_1, X_2 be two orthonormal tangent vectors at ρ and $\xi_{\Phi_1}, \xi_{\Phi_2}$ be their corresponding right-invariant horizontal lifts on the group. If O'Neill's formula can be applied, the sectional curvature of $\mathrm{Dens}(M)$ at point ρ is given by,*

$$K(\rho)(X_1, X_2) \quad = \quad \int_\Omega k(x,1)(\xi_1(x), \xi_2(x)) w(\xi_1(x), \xi_2(x)) \rho(x)\, \mathrm{d}\nu(x) \quad + \quad \frac{3}{4} \left\| [\xi_1, \xi_2]^V \right\|^2 \tag{2.31}$$

where

$$w(\xi_1(x), \xi_2(x)) = g_{\mathcal{C}(M)}(x,1)(\xi_1(x), \xi_1(x)) g_{\mathcal{C}(M)}(x)(\xi_2(x), \xi_2(x)) - \left(g_{\mathcal{C}(M)}(x,1)(\xi_1(x), \xi_2(x))\right)^2$$

and $[\xi_{\Phi_1}, \xi_{\Phi_2}]^V$ denotes the vertical projection of $[\xi_{\Phi_1}, \xi_{\Phi_2}]$ at identity, $\|\cdot\|$ denotes the norm at identity and $k(x,1)$ is the sectional curvature of the cone at point $(x,1)$ in the directions $(\xi_1(x), \xi_2(x))$.

This computation is only formal and we will not attempt here to give a rigorous meaning to this formula similarly to what has been done in [40] for the L^2 Wasserstein metric. Yet, it has interesting consequences: the curvature of the space of densities endowed with the WF metric is always greater or equal than the curvature of the cone $\mathcal{C}(M)$. In particular, it is non-negative if the curvature of (M,g) is bigger than 1, as a consequence of Proposition 4.

3. The corresponding Monge and Kantorovich formulations and the associated polar decomposition

By the geometric point of view developed above, it is possible to derive a Monge formulation directly and also to derive a pre-formulation of the Monge-Ampère equation. We first derive formally the equations for which a precise meaning will be given in the next sections.

3.1. The Monge formulation.

The first important consequence of the L^2 metric on the group and the Riemannian submersion is that one can define a Monge formulation of the Wasserstein-Fisher-Rao metric as follows:

$$\mathrm{WF}(\rho_0, \rho_1) = \inf_{(\varphi,\lambda)} \left\{ \|(\varphi, \lambda) - (\mathrm{Id}, 1)\|_{L^2(\rho_0)} \,:\, \varphi_*(\lambda^2 \rho_0) = \rho_1 \right\} . \tag{3.1}$$

It is then possible to derive a pre-formulation to the classical Monge-Ampère equation at least formally: Under the assumption that there exists a smooth minimizer (φ, λ) of (3.1), there exists a function $p \in C^\infty(M,\mathbb{R})$ given by the Lagrange multiplier rule such that

$$(\varphi(x), \lambda(x)) = \exp_{(x,1)}^{\mathcal{C}(M)} \left(-\frac{1}{2}\nabla p(x), -p(x) \right) , \tag{3.2}$$

where $\exp_{(x,1)}$ denotes the Riemannian exponential map on $\mathcal{C}(M)$. The "pushforward" constraint now reads

$$(|1 - p(x)|^2 + \frac{1}{4}\|\nabla p(x)\|^2)\rho_0(x) = \mathrm{Jac}(\varphi)(x)\, \rho_1(\varphi(x)) \tag{3.3}$$

where

$$\varphi(x) = \exp_x^M \left(-\arctan\left(\frac{\|\nabla p(x)\|}{2(1+p(x))} \right) \frac{\nabla p(x)}{\|\nabla p(x)\|} \right) .$$

Under the change of variable $z \overset{\text{def.}}{=} -\log(1-p)$, the previous equations become

$$(1 + \frac{1}{4}\|\nabla z\|^2)e^{-2z}\rho_0 = \det(D\varphi)\rho_1 \circ \varphi \tag{3.4}$$

and

$$\varphi(x) = \exp_x^M \left(-\arctan\left(\frac{1}{2}\|\nabla z\| \right) \frac{\nabla z(x)}{\|\nabla z(x)\|} \right) .$$

However, the function p (or z) is not completely characterized. Indeed, in standard optimal transport, the optimal potential is convex. Convexity will be replaced by c-concavity for a particular cost. Note that this result has been established in the Euclidean case in [38, Theorem 6.7] under mild assumptions on the densities. Their result is based on a detailed study of the equivalent Kantorovich formulation which is presented in 3.2. This equivalent formulation is only proven in [38] and [12] in the Euclidean case and it is expected to be true [38, Section 8.5] in the Riemannian case.

In the next section, we present the Kantorovich formulation associated with the Monge formulation and we prove the equality between the Kantorovich and the dynamic formulations of the WF metric.

3.2. **The Kantorovich Formulation.** From a variational point of view, it is important to derive a relaxation of the Monge formulation. It is of interest to understand first the simple situation when the source and target measures are single Dirac masses and when M is a convex and compact domain in the Euclidean space as studied in [13]. This also applies to the case of a Riemannian manifold since it can be shown using the static formulation proven in this section.

Proposition 18. *Let M be a convex and compact domain in $\mathbb{R}^d$ with the Euclidean metric. Let $m_1\delta_{x_1}$ and $m_2\delta_{x_2}$ be two Dirac masses with $x_1, x_2 \in M$ and $m_1, m_2 \in \mathbb{R}_+^*$.*

If $d(x_1, x_2) < \pi/2$, there exists a unique geodesic which is $m(t)\delta_{x(t)}$ where $(x(t), m(t))$ is the geodesic in $M \times \mathbb{R}_+^*$ with the cone metric between (x_1, m_1) and (x_2, m_2).

If $d(x_1, x_2) > \pi/2$, there exists a unique geodesic which is $m_1(t)\delta_{x_1} + m_2(t)\delta_{x_2}$ where $m_1(t) = m_1(1-t)^2$ and $m_2(t) = m_2 t^2$ describe the geodesics between (x_i, m_i) and the cone point for $i = 1, 2$.

If $d(x_1, x_2) = \pi/2$, there exists an infinite number of geodesics which are convex combinations of the two first types defined above.

The important point is that passing to the case of measures the angle of the cone has been divided by 2. This is because the optimization problem is not formulated on the space of geodesics on $M \times \mathbb{R}_+^*$, but on the space of measures on M. And, in particular, the cost between Dirac masses has to be convex.

The generalization to any positive Radon measures of the Kantorovich relaxation requires the definition of a convex functional which is one-homogeneous on the space of Radon measures described below. The next theorem is proven in [12] and in another form in [38], both only in the Euclidean case. We now extend it to in a Riemannian setting.

Theorem 19. *For two given positive Radon measures ρ_1, ρ_2, we define, for $\mathcal{M}_+(M^2)$ the space of positive Radon measures on M^2,*

$$\Gamma(\rho_1, \rho_2) \overset{\text{def.}}{=} \left\{ (\gamma_1, \gamma_2) \in \left(\mathcal{M}_+(M^2)\right)^2 : p^1_*\gamma_1 = \rho_1,\, p^2_*\gamma_2 = \rho_2 \right\} , \tag{3.5}$$

where p^1 and p^2 denote the projection on the first and second factors of the product M^2. The variational problem associated with the Wasserstein-Fisher-Rao distance is

$$\mathrm{WF}^2(\rho_1, \rho_2) = \min_{(\gamma_1,\gamma_2)\in\Gamma(\rho_1,\rho_2)} \int_{M^2} d^2_{\mathcal{C}(M)}\left((x, \frac{\mathrm{d}\gamma_1}{\mathrm{d}\gamma}), (y, \frac{\mathrm{d}\gamma_2}{\mathrm{d}\gamma}) \right) \mathrm{d}\gamma(x, y) , \tag{3.6}$$

where $d^2_{\mathcal{C}(M)}$ is the square of the cone distance given in definition 2 and γ is any measure that dominates ρ_1 and ρ_2.

Proof. The proof is given in Appendix B. □

Remark 3. *The fact that* $S(\gamma_1,\gamma_2) \stackrel{\text{def.}}{=} \int_{M^2} d^2_{\mathcal{C}(M)}\left((x, \frac{\mathrm{d}\gamma_1}{\mathrm{d}\gamma}), (y, \frac{\mathrm{d}\gamma_2}{\mathrm{d}\gamma})\right) \mathrm{d}\gamma(x,y)$ *is well defined follows from the application of* [54, Theorem 5]. *It does not depend on the choice of the measure* γ *since the function* d^2 *is one-homogeneous w.r.t. the mass variables. As a consequence of Rockafellar's theorem* [54, Theorem 5], S *is convex and lower-semicontinuous on the space of Radon measures as the Legendre-Fenchel transform of a convex functional on the space of continuous functions.*

We also state without proof the dual formulation which is given by the application of Fenchel-Rockafellar duality theorem, see [13].

Proposition 20. *It holds*

$$\mathrm{WF}^2(\rho_0,\rho_1) = \sup_{(\phi,\psi)\in C(M)^2} \int_M \phi(x)\,\mathrm{d}\rho_0(x) + \int_M \psi(y)\,\mathrm{d}\rho_1(y) \tag{3.7}$$

subject to $\forall (x,y) \in M^2$,

$$\begin{cases} \phi(x) \leq 1\,, \quad \psi(y) \leq 1\,, \\ (1-\phi(x))(1-\psi(y)) \geq \cos^2\left(d(x,y) \wedge (\pi/2)\right)\,. \end{cases} \tag{3.8}$$

A reformulation of this linear optimization problem is

$$\mathrm{WF}^2(\rho_0,\rho_1) = \sup_{(z_0,z_1)\in C(M)^2} \int_M 1 - e^{-z_0(x)}\,\mathrm{d}\rho_0(x) + \int_M 1 - e^{-z_1(y)}\,\mathrm{d}\rho_1(y) \tag{3.9}$$

subject to $\forall (x,y) \in M^2$,

$$z_0(x) + z_1(y) \leq -\log\left(\cos^2\left(d(x,y) \wedge (\pi/2)\right)\right)\,. \tag{3.10}$$

Interestingly, the last formulation can be further reduced since the exponential $r \mapsto e^r$ is the Fenchel-Legendre conjugate associated with the Kullback-Leibler divergence defined below. Therefore, using duality again, it is proven in [38] that the static problem in Proposition 20 can be rewritten as

$$\begin{aligned} \mathrm{WF}^2(\rho_0,\rho_1) = \inf_{\gamma\in\mathcal{M}_+(M^2)} &\mathrm{KL}(\mathrm{Proj}^1_*\gamma,\rho_0) + \mathrm{KL}(\mathrm{Proj}^2_*\gamma,\rho_1) \\ &- \int_{M^2} \log(\cos^2(d(x,y)\wedge(\pi/2)))\,\mathrm{d}\gamma(x,y) \end{aligned} \tag{3.11}$$

with

$$\mathrm{KL}(\mu,\nu) = \int \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right) \mathrm{d}\nu + |\nu| - |\mu|\,,$$

the Kullback-Leibler divergence. Formulation (3.11) of unbalanced optimal transport and its extensions have been intensively developed in [38], where generalizations of this metric are studied in spaces such as Hausdorff topological spaces endowed with a (pseudo) distance satisfying mild conditions. More interestingly, in our situation where the underlying space M is a finite dimensional Riemannian manifold, is the existence of solutions to the dual problem (3.9) - (3.12), which is proven in [38]. In particular, we rely on the existence results given in [38, Section 6.2] in a larger class of potentials than simply continuous. We also strongly rely on their characterization presented in [38, Theorem 6.3].

3.3. **Polar factorization on the automorphism group of half-densities and Monge-Ampère equation.** One can state a similar version to Brenier's polar factorization theorem [5] in this context. Actually, this can be understood as a constrained version of the polar factorization, since the diffeomorphisms are restricted to be automorphisms. To the best of our knowledge, this is not a direct consequence of [42] or [38]. However, adapting arguments from these two articles, under mild conditions on the initial and final densities, we give a rigorous meaning to the equivalent of the Monge-Ampère equation and we propose a polar factorization on the automorphism group of the cone.

We use define a c-concave function as in [42].

Definition 6. A function $z\colon M \to \mathbb{R}\cup\{\pm\infty\}$ is said to be c-concave if there exists $\bar{z}\colon M \to \mathbb{R}\cup\{\pm\infty\}$ such that

$$z(x) = \inf_{y\in M} \left[c(x,y) - \bar{z}(y)\right], \quad \forall x \in M. \tag{3.12}$$

Definition 7 (Admissible measures)**.** We say that a positive Radon measure ρ on M is admissible (with respect to vol) if for any $x \in M$, there exists $y \in \mathrm{Supp}(\rho)$ such that $d(x,y) < \pi/2$.

Note that when the diameter of M is less than $\pi/2$ then all but the null measure are admissible. In general, when the null measure is considered, the geodesic is unique and is the geodesic corresponding to the Hellinger distance. To shorten the notations, we denote by c the function $c(x,y) = -\log\left(\cos^2\left(d(x,y)\wedge(\pi/2)\right)\right)$ which intervenes in the constraint (3.12) of the dual formulation. We prove that a solution of WF, in the form (3.9) - (3.12) leads to a solution of the associated Monge problem defined by (3.1).

Lemma 21 (sub-differentiability)**.** *Let $y \in M$, the function g defined on M by $g(x) = \cos^2\left(d(x,y)\right)$ is sub-differentiable.*

Proof. The function $d^2(\cdot,y)$ is super-differentiable see [42, Proposition 6] for instance. Therefore $d^2_{\pi/2}(\cdot,y) = (d(x,y)\wedge(\pi/2))$ is also super-differentiable and the function g is sub-differentiable as the combinaison of a decreasing C^1 function and the super-differentiable function $d^2_{\pi/2}(\cdot,y)$, see [42, Lemma 5]. $\square$

Proposition 22 (Approximate differentiability and optimality)**.** *Let ρ_0, ρ_1 be two positive Radon measures and (z_0,z_1) be the generalized optimal potentials for $\mathrm{WF}^2(\rho_0,\rho_1)$. Suppose that ρ_0 and ρ_1 are admissible and $\rho_0 << \mathrm{vol}$, then z_0 is ρ_0 a.e. unique and approximate differentiable on Supp(ρ_0). The optimal plan γ in the formulation (3.11) is unique, with marginals $\gamma_0 = e^{-z_0}\rho_0$, $\gamma_1 = e^{-z_1}\rho_1$ and concentrated on the graph of*

$$x \mapsto \varphi(x) = \exp_x^M\left(-\arctan\left(\frac{\|\tilde{\nabla} z_0(\bar{x})\|}{2}\right)\frac{\tilde{\nabla} z_0(\bar{x})}{\|\tilde{\nabla} z_0(\bar{x})\|}\right), \tag{3.13}$$

that is $\varphi_*\gamma_0 = \gamma_1$ and $\gamma = (\mathrm{Id}\times\varphi)_*\gamma_0$. Finally

$$\mathrm{WF}^2(\rho_0,\rho_1) = \int_M 1 - e^{-z_0(x)}\,\mathrm{d}\rho_0(x) + \int_M 1 - e^{-z_1(y)}\,\mathrm{d}\rho_1(y)\,. \tag{3.14}$$

Note that (z_0,z_1) may not be admissible in (3.9) but (3.14) still holds true. The proof of this proposition (being more technical) is given in Appendix B, we prefer to discuss the corresponding formulation of the Monge-Ampère equation hereafter.

Following Brenier's approach in the case of optimal transport, see [5, Section 1.4] and [57, Section 12], we expect the potential found in Proposition 22, denoted by z, to be a solution of a Monge-Ampère equation. To formally derive the equation we suppose that z is smooth. Recall that $c(x,y) = -\log(\cos^2(d_{\pi/2}(x,y)))$ and $\varphi(x) = \exp_x^M\left(-\arctan\left(\frac{1}{2}\|\nabla z(x)\|\right)\frac{\nabla z(x)}{\|\nabla z(x)\|}\right)$, therefore

$$2\sqrt{2}\tan(d_{\pi/2}(x,y))\frac{\sqrt{2}}{2d_{\pi/2}(\bar{x},\bar{y})}\nabla\left(\frac{1}{2}d^2_{\pi/2}(\bar{x},\bar{y})\right) = (\nabla_x c)(x,\varphi(x))$$

and the sub-differentiable equality (B.14) reads

$$\nabla z(x) - (\nabla_x c)(x,\varphi(x)) = 0\,. \tag{3.15}$$

Differentiating (3.15) and taking the determinant yields

$$\det\left[-\nabla^2 z(x) + (\nabla^2_{xx}c)(x,\varphi(x))\right] = \left|\det\left[(\nabla_{x,y}c)(x,\varphi(x))\right]\right|\left|\det(\nabla\varphi)\right|. \tag{3.16}$$

Notice that the c-convexity property of z implies that $-\nabla^2 z + (\nabla^2_{xx} c)(x, \varphi(x))$ is a nonnegative symmetric matrice. To obtain the equation on z, we observe that $\varphi_* \left((1 + \frac{1}{4}\|\nabla z\|^2)e^{-2z}\rho_0\right) = \rho_1$ (see the proof of Proposition 23 below for details) or equivalently

$$|\det(\nabla\varphi)| = e^{-2z}\left(1 + \frac{1}{4}\|\nabla z\|^2\right)\frac{f}{g \circ \varphi}\,,$$

for smooth z and smooth positive measures ρ_0 and ρ_1 with densities f and g with respect to the volume measure vol. Together with (3.16), we obtain the WF-Monge-Ampère equation defined by (3.17)

$$\det\left[-\nabla^2 z(x) + (\nabla^2_{xx} c)(x, \varphi(x))\right] = \left|\det\left[(\nabla_{x,y} c)(x, \varphi(x))\right]\right| e^{-2z(x)}\left(1 + \frac{1}{4}\|\nabla z(x)\|^2\right)\frac{f(x)}{g \circ \varphi(x)}\,,$$

where φ is given by (3.18) and satisfies the second boundary value problem: φ maps the support of ρ_0 to the support of ρ_1. Following Brenier [5, Section 1.4], Proposition 23 below can be taken as a definition of weak solutions for the WF-Monge-Ampère equation with second boundary value problem. The question of regularity for the potential z solution of a WF-Monge-Ampère equation will be studied elsewhere. One would first need to compute the Ma-Trudinger-Wang tensor associated to c see [18], [57, Section 12].

Remark 4. *In optimal transport, the notation* $\text{c-exp}_x(v) = \left[(-\nabla_x c)(x, \cdot)\right]^{-1}(v)$ *is classical. Using this notation one remarks that* (3.15) *implies*

$$\varphi(x) = \text{c-exp}(-\nabla z(x)).$$

Proposition 23 (Brenier's weak solution of WF-Monge-Ampère)**.** *Let ρ_0, ρ_1 be two admissible measures such that ρ_0 has density w.r.t. the volume measure on M. Then, there exists a ρ_0 a.e. unique c-convex function on M, z, approximatively differentiable ρ_0-a.e., such that the associated unbalanced transport couple (φ, λ) defined by*

$$\varphi(x) = \exp^M_x\left(-\arctan\left(\frac{1}{2}\|\tilde{\nabla} z(x)\|\right)\frac{\tilde{\nabla} z(x)}{\|\tilde{\nabla} z(x)\|}\right) \tag{3.18}$$

and

$$\lambda(x) = e^{-z(x)}\sqrt{1 + \frac{1}{4}\|\tilde{\nabla} z(x)\|^2} \tag{3.19}$$

satisfies

$$\pi[(\varphi, \lambda), \rho_0] = \varphi_*\left(\lambda^2 \rho_0\right) = \varphi_*\left((1 + \frac{1}{4}\|\tilde{\nabla} z\|^2)e^{-2z}\rho_0\right) = \rho_1\,. \tag{3.20}$$

Moreover, (φ, λ) is the unique ρ_0 a.e. unbalanced transport couple associated to a c-convex potential,also unique, such that $\pi[(\varphi, \lambda), \rho_0] = \rho_1$. The potential z is a weak solution for the WF*-Monge-Ampère equation* (3.17) *with second boundary value problem and is characterized by*

$$\mathrm{WF}^2(\rho_0, \rho_1) = \int_M 1 - e^{-z(x)}\,\mathrm{d}\rho_0(x) + \int_M 1 - e^{-z^c(y)}\,\mathrm{d}\rho_1(y)\,. \tag{3.21}$$

Proof. Let (z_0, z_1) be the optimal potentials for $\mathrm{WF}^2(\rho_0, \rho_1)$. From Proposition 22, we know that $x \mapsto \varphi(x) = \exp^M_x\left(-\arctan\left(\frac{\|\tilde{\nabla} z_0(x)\|}{2}\right)\frac{\tilde{\nabla} z_0(x)}{\|\tilde{\nabla} z_0(x)\|}\right)$ is well defined ρ_0 a.e. and $\varphi_*(\gamma_0) = \gamma_1$ where

$\gamma_i = \sigma_i\rho_i = e^{-z_i}\rho_i$, $i = 0,1$. Therefore

$$\begin{aligned}\rho_1 &= \sigma_1^{-1}\gamma_1 = \sigma_1^{-1}\varphi_*(\gamma_0) = \sigma_1^{-1}\varphi_*\left(\sigma_0\rho_0\right)\\ &= \varphi_*\left(e^{-z_0}\sigma_1^{-1}\circ\varphi\rho_0\right) = \varphi_*\left(e^{-z_0}e^{z_1\circ\varphi}\rho_0\right) = \varphi_*\left(e^{-z_0}e^{c(\cdot,\varphi(\cdot))}e^{-z_0}\rho_0\right)\\ &= \varphi_*\left(e^{-2z_0}\left(1+\frac{1}{4}\|\tilde\nabla z_0\|^2\right)\rho_0\right) = \varphi_*\left(\left(e^{-z_0}\sqrt{1+\frac{1}{4}\|\tilde\nabla z_0\|^2}\right)^2\rho_0\right)\\ &= \pi\left[\left(\varphi, e^{-z_0}\sqrt{1+\frac{1}{4}\|\tilde\nabla z_0\|^2}\right),\rho_0\right].\end{aligned}$$

We used that ρ_0 a.e. $z_0(x)+z_1(\varphi(x)) = c(x,\varphi(x))$, $1+\tan^2(x) = 1/\cos^2(x)$ and thus $1+\frac{1}{4}\|\tilde\nabla z_0(x)\|^2 = e^{c(x,\varphi(x))}$. Equation (3.14) is exactly (3.21).

To prove uniqueness, consider z to be a c-convex function, such that (φ,λ) are well defined through (3.18) and (3.20) and $\pi[(\varphi,\lambda),\rho_0] = \rho_1$. Then, we claim that $\gamma = [\mathrm{Id}\times\varphi]_*(e^{-z}\rho_0)$ is an optimal plan for $\mathrm{WF}^2(\rho_0,\rho_1)$ in (3.11). Indeed, let us check that γ satisfies the optimality conditions of [38, Theorem 6.3(b)].

- γ is concentrated on the set of equality for a pair (z,z^c) of c-convex functions. By definition of φ, it holds ρ_0 a.e. and therefore $\gamma_0 = e^{-z}\rho_0$ a.e.

$$z(x) + z^c(\varphi(x)) = c(x,\varphi(x))\,. \tag{3.22}$$

  Thus, (z,z^c) satisfies for all $(x,y)\in M\times M$, $z(x)+z^c(y)\leq c(x,y)$ with equality γ a.e.
- The marginals are absolutely continuos with respect to ρ_0 and ρ_1. It holds true for $\gamma_0 = e^{-z}\rho_0$. Notice then that ρ_0 a.e.

$$\lambda^2(x) = e^{-2z(x)}(1+\frac{1}{4}\|\tilde\nabla z(x)\|^2) = e^{-z(x)}e^{z^c(\varphi(x))}\,.$$

  It yields

$$\rho_1 = \varphi_*(\lambda^2\rho_0) = \varphi_*(e^{z^c(\varphi(x))}e^{-z(x)}\rho_0) = e^{z^c}\varphi_*(\gamma_0) = e^{z^c}\gamma_1\,,$$

  thus $\gamma_1 = e^{-z^c}\rho_1$ and γ is optimal for $\mathrm{WF}^2(\rho_0,\rho_1)$.

The computation (B.15) yields (3.21) and the uniqueness of the generalized optimal potentials for $\mathrm{WF}^2(\rho_0,\rho_1)$ in Proposition (22) implies the uniqueness of (z,φ,λ).

□

Note that if z is smooth, it satisfies (3.17). It turns out that this factorization can be extended to a larger class containing the automorphism group of the cone $\mathrm{Aut}(\mathcal{C}(M))$. In the following, we state a polar factorization theorem for a class of maps from M to $C(M)$. We start with definitions.

Definition 8. We define the generalized automorphism semigroup of $C(M)$ as the set of mesurable maps (φ,λ) from M to $C(M)$

$$\overline{\mathrm{Aut}}(\mathcal{C}(M)) = \left\{(\varphi,\lambda)\in\mathcal{M}es(M,M)\ltimes\mathcal{M}es(M,\mathbb{R}_+^*)\right\}\,, \tag{3.23}$$

endowed with the semigroup law

$$(\varphi_1,\lambda_1)\cdot(\varphi_2,\lambda_2) = (\varphi_1\circ\varphi_2,(\lambda_1\circ\varphi_2)\lambda_2)\,.$$

We also consider the stabilizer of the volume measure in the automorphisms of $C(M)$. It is a subsemigroup and is defined by

$$\overline{\mathrm{Aut}}_{\mathrm{vol}}(\mathcal{C}(M)) = \left\{(s,\lambda)\in\overline{\mathrm{Aut}}(\mathcal{C}(M))\,:\,\pi\left((s,\lambda),\mathrm{vol}\right) = \mathrm{vol}\right\}\,. \tag{3.24}$$

By abuse of notation, any $(s,\lambda)\in\overline{\mathrm{Aut}}_{\mathrm{vol}}(\mathcal{C}(M))$ will be denoted $\left(s,\sqrt{\mathrm{Jac}(s)}\right)$ meaning that for every continuous function $f\in C(M,\mathbb{R})$

$$\int_M f(s(x))\sqrt{\mathrm{Jac}(s)}^2\,\mathrm{d}\,\mathrm{vol}(x) = \int_M f(x)\,\mathrm{d}\,\mathrm{vol}(x)\,. \tag{3.25}$$

Theorem 24 (Polar factorization)**.** *Let $(\phi, \lambda) \in \overline{\mathrm{Aut}}(\mathcal{C}(M))$ be an element of the generalized automorphism group of the half-densities bundle such that $\rho_1 = \pi_0\left[(\phi,\lambda), \mathrm{vol}\right]$ is an absolute continuous admissible measure. Then, there exists a unique minimizer, characterized by a c-convex function z_0, to the Monge formulation (3.1) between* vol *and ρ_1 and there exists a unique measure preserving generalized automorphism $(s, \sqrt{\mathrm{Jac}(s)}) \in \overline{\mathrm{Aut}}_{\mathrm{vol}}(\mathcal{C}(M))$ such that* vol *a.e.*

$$(\phi,\lambda) = \exp^{\mathcal{C}(M)}\left(-\frac{1}{2}\nabla p_{z_0}, -p_{z_0}\right) \circ (s, \sqrt{\mathrm{Jac}(s)}) \tag{3.26}$$

or equivalently

$$(\phi,\lambda) = \left(\varphi, e^{-z_0}\sqrt{1+\|\nabla z_0\|^2}\right) \cdot (s, \sqrt{\mathrm{Jac}(s)})\,, \tag{3.27}$$

where $p_{z_0} = e^{z_0} - 1$ and

$$\varphi(x) = \exp_x^M\left(-\arctan\left(\frac{1}{2}\|\nabla z_0(x)\|\right)\frac{\nabla z_0(x)}{\|\nabla z_0(x)\|}\right)\,. \tag{3.28}$$

Moreover $(s, \sqrt{\mathrm{Jac}(s)})$ is the unique $L^2(M, \mathcal{C}(M))$ projection of (ϕ,λ) onto $\overline{\mathrm{Aut}}_{\mathrm{vol}}(\mathcal{C}(M))$.

We also state a more intrinsic formulation of the theorem.

Corollary 25. *Denote by $\mathcal{M}es^1(\mathcal{C}(M)))^{\mathbb{R}_+^*}$ the space of mesurable and approximate differentiable functions $f : \mathcal{C}(M) \mapsto \mathbb{R}$ that satisfy $f(x,r) = r^2 f(x,1)$ for any $r \in \mathbb{R}_+^*$. Under the hypothesis of Theorem 23, there exists a unique couple $\left((s, \sqrt{\mathrm{Jac}(s)}), \Psi_p\right) \in \overline{\mathrm{Aut}}_{\mathrm{vol}} \times \mathcal{M}es^1(\mathcal{C}(M)))^{\mathbb{R}_+^*}$ such that*

$$(\phi,\lambda) = \exp^{\mathcal{C}(M)}(-\nabla\Psi_p) \circ (s, \sqrt{\mathrm{Jac}(s)})\,, \tag{3.29}$$

where $\Psi_p(x,r) = r^2 p(x)$.

Proof of Theorem 24. We denote $\rho_0 = \mathrm{vol}$ and $\rho_1 = \pi_0\left[(\phi,\lambda), \rho_0\right]$. Let (z_0, z_1) be a solution of $\mathrm{WF}^2(\rho_0,\rho_1)$ and γ be an optimal unbalanced transport plan. By symmetry, (z_1, z_0) is a solution of $\mathrm{WF}^2(\rho_1,\rho_0)$ and γ^t is an optimal unbalanced transport plan. Let finally (φ_0,λ_0) and (φ_1,λ_1) be the two transport couples given by application of Proposition 22 to (ρ_0,ρ_1) and (ρ_1,ρ_0). We divide the proof into four small steps. We also denote $\mathrm{dom}(f)$ the domain of definition of the function f.

Step 1: φ_0 and φ_1 are inverse maps. On $U = \varphi_0^{-1}(\mathrm{dom}\,\tilde\nabla z_1) \cap \mathrm{dom}(\tilde\nabla z_0)$ which has full γ_0 and therefore ρ_0 measure (we use here the admissible condition to say that γ_0 and ρ_0 have the same support), we have

$$z_0(x) + z_1(\varphi_0(x)) = c(x, \varphi_0(x))$$

and thus $\varphi_1(\varphi_0(x)) = x$. Similarly, it holds $\varphi_0(\varphi_1(y)) = y$ on $V = \varphi_1^{-1}(\mathrm{dom}\,\nabla z_0) \cap \mathrm{dom}(\nabla z_1)$ which has full ρ_1 measure.

Step 2: (φ_0,λ_0) and (φ_1,λ_1) are inverse in $\overline{\mathrm{Aut}}$. From Step 1, ρ_1 a.e. it holds $\varphi_0(\varphi_1(y)) = y$. Thus, ρ_1 a.e.

$$(\varphi_0,\lambda_0)\cdot(\varphi_1,\lambda_1) = (\varphi_0\circ\varphi_1, \lambda_0\circ\varphi_1\lambda_1) = (\mathrm{Id}, (\lambda_0\circ\varphi_1)\lambda_1)\,.$$

Moreover by (3.20) of Proposition 23 applied twice

$$\pi\left[(\varphi_0,\lambda_0)\cdot(\varphi_1,\lambda_1), \rho_1\right] = \pi\left[(\varphi_0,\lambda_0), \pi\left[(\varphi_1,\lambda_1),\rho_1\right]\right] = \pi\left[(\varphi_0,\lambda_0),\rho_0\right] = \rho_1\,.$$

It implies that

$$\pi\left[(\mathrm{Id}, (\lambda_0\circ\varphi_1)\lambda_1), \rho_1\right] = \pi\left[(\varphi_0,\lambda_0)\cdot(\varphi_1,\lambda_1),\rho_1\right] = \rho_1\,.$$

In other words, we have ρ_1 a.e. $(\lambda_0\circ\varphi_1)\lambda_1 = 1$ and ρ_1 a.e.

$$(\varphi_0,\lambda_0)\cdot(\varphi_1,\lambda_1) = (\mathrm{Id}, 1)\,.$$

Step 3: polar factorization. Let $(s,\lambda_s) = (\varphi_1,\lambda_1)\cdot(\phi,\lambda) = (\varphi_1\circ\phi, \lambda_1\circ\phi\lambda)$. By construction, one has

$$\pi\left[(s,\lambda_s),\rho_0\right] = \pi\left[(\varphi_1,\lambda_1)\cdot(\phi,\lambda),\rho_0\right] = \pi\left[(\varphi_1,\lambda_1), \pi\left[(\phi,\lambda),\rho_0\right]\right] = \pi\left[(\varphi_1,\lambda_1),\rho_1\right] = \rho_0\,.$$

Therefore, (s,λ_s) belongs to $\overline{\mathrm{Aut}}_{\mathrm{vol}}$ and $\lambda_s = \sqrt{\mathrm{Jac}(s)}$ holds in the weak sense (3.25). Thus

$$(\phi,\lambda) = (\mathrm{Id},1)\cdot(\phi,\lambda) = (\varphi_0,\lambda_0)\cdot(\varphi_1,\lambda_1)\cdot(\phi,\lambda) = (\varphi_0,\lambda_0)\cdot(s,\sqrt{\mathrm{Jac}(s)})\,.$$

It proves the polar factorization.

Step 4: Uniqueness. The pair of c-convex potentials (z_0,z_1) is optimal for $\mathrm{WF}(\rho_0,[(\varphi_0,\lambda_0),\rho_0]) = \mathrm{WF}(\rho_0,\rho_1)$ and therefore by Proposition 23, z_i are unique ρ_i a.e.. We deduce that the projection $(s,\sqrt{\mathrm{Jac}(s)}) = (\varphi_1,\lambda_1)\cdot(\phi,\lambda)$ is also unique ρ_0 a.e.. Indeed the positivity of λ implies that $\mathrm{Supp}(\lambda^2\rho_0) = \mathrm{Supp}(\rho_0)$, thus ϕ maps $\mathrm{Supp}(\rho_0)$ onto $\mathrm{Supp}(\rho_1)$ and the uniqueness of φ_1 and λ_1, ρ_1 a.e., implies the uniqueness of s and $\sqrt{\mathrm{Jac}(s)}$, ρ_0 a.e.. To prove that $(s,\sqrt{\mathrm{Jac}(s)})$ is the $L^2(M,\mathcal{C}(M))$ projection of (ϕ,λ) onto $\overline{\mathrm{Aut}}_{\mathrm{vol}}(\mathcal{C}(M))$, we observe

$$\begin{aligned}\inf_{(\sigma,\sqrt{\mathrm{Jac}(\sigma)})\in\overline{\mathrm{Aut}}_{\mathrm{vol}}(\mathcal{C}(M))} \left\|(\phi,\lambda)-(\sigma,\sqrt{\mathrm{Jac}(\sigma)})\right\|^2_{L^2(\rho_0)} &\geq \mathrm{WF}^2(\rho_0,\rho_1)\\ &= \|(\varphi_0,\lambda_0)-(\mathrm{Id},1)\|^2_{L^2(\rho_0)}\\ &= \left\|(\varphi_0,\lambda_0)\cdot(s,\sqrt{\mathrm{Jac}(s)})-(s,\sqrt{\mathrm{Jac}(s)})\right\|^2_{L^2(\rho_0)} = \left\|(\phi,\lambda)-(s,\sqrt{\mathrm{Jac}(s)})\right\|^2_{L^2(\rho_0)}\,,\end{aligned}$$

which gives the result.

□

This polar factorization could probably be understood, under minor verifications, in the framework of abstract polar factorization of Brenier [5]. As in the case of classical optimal transport, Theorem 24 could be extended, for example, to any admissible ρ_1 without the absolute continuity assumption. In this case, one looses uniqueness of the measure preserving generalized automorphism $(s,\sqrt{\mathrm{Jac}(s)})$. An other direction is to project on the subset of $\overline{\mathrm{Aut}}(\mathcal{C}(M))$:

$$\overline{\mathrm{Aut}}_{\rho_0,\mu_0}(\mathcal{C}(M)) = \left\{(s,\lambda)\in\overline{\mathrm{Aut}}(\mathcal{C}(M))\,\middle|\,\pi\left((s,\lambda),\rho_0\right)=\mu_0\right\}\,,$$

in the spirit of [56, Theorem 3.15]. The proof is similar to the one given above. This polar factorization also yields by linearization an Helmholtz decomposition of velocity vector fields. We will not go further in these directions and leave it for future works.

4. The Euler-Arnold equation and the H^{div} right-invariant metric on the diffeomorphism group

A prototypical example of the situation we are interested in is the case of the incompressible Euler equation. As shown by Arnold [2], the incompressible Euler equation is the Euler-Lagrange equation of geodesics on the group of volume preserving diffeomorphisms for the L^2 right-invariant metric. Let us motivate this section with the following simple proposition whose proof is omitted.

Proposition 26. *Consider a Riemannian submersion constructed as in Proposition 10. Let H_0 be the isotropy subgroup of ρ_0, then, considering H_0 as a Riemannian submanifold of H and denoting G_{H_0} its induced metric, G_{H_0} is a right-invariant metric on H_0.*

It is therefore interesting to start with this point of view, a right-invariant metric on a group of diffeomorphisms and to write down the corresponding geodesic equations. The right-invariance implies that the geodesic equation can be written on the Lie algebra or the tangent space at identity ($T_{\mathrm{Id}}G$ for a Lie group G). This is the case of the usual formulation incompressible Euler equation as in Equation 4.11 and this is the point of view taken in [2]. Actually, it is a particular case of Lagrangians that can be written by a change of variable only at the tangent space of identity $\mathfrak{g} \overset{\text{def.}}{=} T_{\mathrm{Id}}G$, the Lie algebra under the constraint of the flow equation. This class of Lagrangians leads to the so-called Euler-Poincaré or Euler-Arnold equation when the Euler-Lagrange equation is written on $T_{\mathrm{Id}}G$. We describe the derivation of this Euler-Lagrange equation in the next paragraph.

4.1. **The Euler-Arnold equation.** A short proof of the derivation of this equation is given in [49, Theorem 3.2] in the case of a kinetic energy but let us underline that the same equation holds true for general Lagrangians that are right-invariant. We will need the definition of the adjoint and co-adjoint operators:

Definition 9. Let G be a Lie group and $h \in G$, the adjoint operator $\mathrm{Ad}_h : G \times \mathfrak{g} \mapsto \mathfrak{g}$ is defined by

$$\mathrm{Ad}_h(v) \overset{\text{def.}}{=} dL_h \cdot dR_{h^{-1}}(v)\,. \tag{4.1}$$

Then, Ad_h^* is the adjoint of Ad_h defined by duality on $\mathfrak{g}$.
Their corresponding differential map at Id are respectively denoted by ad and ad^*.

Let G be a Lie group, and $\mathcal{L} : TG \mapsto \mathbb{R}$ be a Lagrangian which satisfies the following property,

$$\mathcal{L}(g, \dot{g}) = \mathcal{L}(\mathrm{Id}, dR_{g^{-1}}(\dot{g}))\,. \tag{4.2}$$

The reduced Lagrangian is $\ell : \mathfrak{g} \mapsto \mathbb{R}$ defined by $\ell(v) = \mathcal{L}(\mathrm{Id}, v)$ for $v \in \mathfrak{g}$.
Thus, the variational problem for a reduced Lagrangian reads

$$\inf \int_0^1 \ell(v)\,\mathrm{d}t \quad \text{subject to} \quad \begin{cases} \dot{g} = dR_g(v) \\ g(0) = g_0 \in G \text{ and } g(1) = g_1 \in G\,. \end{cases} \tag{4.3}$$

In order to compute the Euler-Lagrange equation for (4.3), one needs to compute the variation of v in terms of the variation of g. It is given by $\dot{w} - \mathrm{ad}_v\, w$ for any path $w(t) \in T_{\mathrm{Id}}G$, therefore, the Euler-Lagrange equation reads

$$(\partial_t + \mathrm{ad}_v^*)\,\frac{\partial \ell}{\partial v} = 0\,. \tag{4.4}$$

4.2. **The particular case of H^{div} and the Camassa-Holm equation.** When the Lagrangian is a kinetic energy, $\ell(v) = \frac{1}{2}\langle v, Lv\rangle$, which will be also denoted by $\frac{1}{2}\|v\|_{\mathfrak{g}}^2$, where $L : \mathfrak{g} \mapsto \mathfrak{g}$ is a quadratic form and $\langle \cdot, \cdot \rangle$ denotes the dual pairing, one has $\frac{\delta \ell}{\delta v} = Lv$ and Lv is the so-called momentum. Then, the critical curves are determined by their initial conditions $(g(0), \dot{g}(0))$ and the Euler-Poincaré equation (4.4). In the context of infinite dimensional Riemannian manifolds enjoying a group structure, this equation is called the Euler-Arnold equation. Let us compute more explicitely the Euler-Arnold equation and detail the expression of the adjoint Ad_h^* which acts on 1-forms. Let m be a 1-form density, then $\mathrm{Ad}_\varphi^*(\mathsf{m}) = D\varphi^T(\mathsf{m} \circ \varphi)\,\mathrm{Jac}(\varphi)$ and therefore the differentiation w.r.t. φ gives

$$\mathrm{ad}_u^*(\mathsf{m}) = \mathrm{div}(u)\mathsf{m} + Du^T \cdot \mathsf{m} + D\mathsf{m} \cdot u\,. \tag{4.5}$$

Thus, the Euler-Arnold equation reads

$$\begin{cases} \partial_t \mathsf{m}_t + \mathrm{div}(u_t)\mathsf{m}_t + Du_t^T \cdot \mathsf{m}_t + D\mathsf{m}_t \cdot u_t = 0 \\ Lu_t = m_t\,, \end{cases} \tag{4.6}$$

where L is the differential operator defining the metric. A more geometrical way of writing this equation is the following,

$$\partial_t \mathsf{m}_t + \mathcal{L}_{u_t}\mathsf{m}_t + \mathrm{div}(u_t)\mathsf{m}_t = 0\,, \tag{4.7}$$

or alternatively

$$(\partial_t + \mathcal{L}_{u_t})\,(\mathsf{m}_t \otimes \mathrm{vol}) = 0\,, \tag{4.8}$$

together with the relation $Lu_t = \mathsf{m}_t$.

Some important examples in fluid dynamics of the Euler-Arnold equation are given hereafter. For the L^2 metric in one dimension, $Lu = u$, one has

$$\partial_t u + 3\partial_x u\, u = 0\,, \tag{4.9}$$

which is the inviscid Burgers equation.
For the H^{div} metric in one dimension, $Lu = u - \partial_{xx} u$, one has the Camassa-Holm equation (actually when $a = b = 1$)

$$a^2 \partial_t u - b^2 \partial_{txx} u + 3a^2 \partial_x u\, u - 2b^2 \partial_{xx} u\, \partial_x u - b^2 \partial_{xxx} u\, u = 0\,. \tag{4.10}$$

The Korteweg-de Vries equation can also be understood in this setting on a central extension of the group $\mathrm{Diff}(S_1)$. In the case where $G = \mathrm{SDiff}(M)$ is the group of volume preserving diffeomorphisms, the Euler-Arnold equation is the incompressible Euler equation

$$\partial_t u + \nabla_u u = -\nabla p\,,\ \mathrm{div}(u) = 0\,. \tag{4.11}$$

Let us detail the case of the $H^{\mathrm{div}}(\mathbb{T}_d)$ where $\mathbb{T}^d \stackrel{\text{def.}}{=} \mathbb{R}^d/\mathbb{Z}^d$ metric which is detailed in [31, Theorem A.1]. The differential operator takes the form $Lu = a^2 u + b^2 \nabla \,\mathrm{div}(u)$ which gives

$$\begin{aligned} \partial_t Lu + a^2 \left(\mathrm{div}(u) u + \frac{1}{2} \nabla \langle u, u\rangle + Du \cdot u \right) + \\ b^2 \Big(\mathrm{div}(u) \nabla \,\mathrm{div}(u) + Du^T \cdot \nabla \,\mathrm{div}(u) + D[\nabla \,\mathrm{div}(u)] \cdot u \Big) = 0\,. \end{aligned} \tag{4.12}$$

On a Riemannian manifold (M, g), this equation can be written as

$$\partial_t Lu + a^2 \left(\mathrm{div}(u) u^\flat + \mathrm{d}\langle u, u\rangle + \iota_u \,\mathrm{d}u^\flat \right) + b^2 \left(\mathrm{div}(u)\, \mathrm{d}\delta u^\flat + \mathrm{d}\iota_u \,\mathrm{d}\delta u^\flat \right) = 0\,. \tag{4.13}$$

where the notation $\flat$ corresponds to lowering the indices. More precisely, if $u \in \chi(M)$ then $u^\flat$ is the 1-form defined by $v \mapsto g(u, v)$. The notation δ is the formal adjoint to the exterior derivative d and ι is the insertion of vector fields which applies to forms.

In Section 6, we rewrite the Camassa-Holm equation (4.10) as an incompressible Euler equation formulated as (4.11).

4.3. **Smoothness of the flow and metric properties.** For the sake of completeness, we recall in this section some previous works concerning the Camassa-Holm equation as a geodesic equation on the group of diffeomorphisms for the H^{div} right-invariant metric. For instance, the reader can refer to [48] or [35] where much more results are proven. Using the Ebin and Marsden approach in [19], the geodesic equation can be interpreted as an ODE on a Hilbert space. For that purpose, one needs to consider the geodesic equation on a sufficiently regular Sobolev space H^s, for $s > d/2 + 2$. The key point consists in switching from Eulerian to Lagrangian coordinates which enables to prove the smoothness of the metric. This is enough if the metric is *strong*, for instance the right-invariant H^s metric, since one can apply general results from Riemannian geometry in infinite dimensions [36]. However, since the H^{div} metric is of H^1 type, a direct proof that the geodesic spray is smooth is needed. Indeed, in this case, the topology defined by the metric is weaker than that of the space in which the geodesic are studied (see [19] for the definition).

Theorem 27 (Ebin and Marsden)**.** *Let M be a compact manifold without boundary. On $\mathrm{Diff}^s(M)$ for $s > d/2+1$, the H^{div} right-invariant metric is a smooth and weak Riemannian metric. Moreover, if $s > d/2 + 2$, the exponential map is smooth and locally defined on $T\,\mathrm{Diff}^s(M)$.*

Although this theorem is not stated in this particular form in [19], this result can be seen as a byproduct of their results as explained in [49, Theorem 4.1].

Remark 5. *Although this theorem is stated in a smooth Sobolev setting, at least H^s for $s > d/2+1$, the result is not trivial since the composition $\mathrm{Diff}^s(M) \times \mathrm{Diff}^s(M) \mapsto \mathrm{Diff}^s(M)$ defined by $(\varphi, \psi) \mapsto \varphi \circ \psi$ is smooth w.r.t. φ (because linear) but it is not smooth w.r.t. ψ. Therefore the fact that the metric defined below in (4.14) is smooth on $\mathrm{Diff}^s(M)$ is not directly given by working in a smooth enough Sobolev setting.*

Proof of the smoothness of the metric. First recall that $H^s(M)$ is embedded in $C^1(M)$ for $s > d/2+1$ and it is a Hilbert algebra if $s > d/2$ which means that the product of two functions is a bounded bilinear operation. The idea consists in writing the H^{div} metric in Lagrangian coordinates. Consider $X_\varphi \in H^s(M)$ a tangent vector at $\varphi \in \mathrm{Diff}^s(M)$, the metric reads

$$G_\varphi(X_\varphi, X_\varphi) = \int_M a^2 |X_\varphi \circ \varphi^{-1}|^2 + b^2 \operatorname{div}(X_\varphi \circ \varphi^{-1})^2 \,\mathrm{dvol}\,. \tag{4.14}$$

Note that the differentiation of the composition can be written as

$$D(X_\varphi \circ \varphi^{-1}) = (DX_\varphi \cdot [D\varphi]^{-1}) \circ \varphi^{-1}\,, \tag{4.15}$$

where the symbol $\cdot$ denotes matrix multiplication. By the change of variable by φ, one has

$$G_\varphi(X_\varphi, X_\varphi) = \int_M a^2 |X_\varphi|^2 \operatorname{Jac}(\varphi) + b^2 \left(\operatorname{Tr}(DX_\varphi \cdot [D\varphi]^{-1})\right)^2 \operatorname{Jac}(\varphi) \,\mathrm{dvol}\,. \tag{4.16}$$

Therefore, the metric only involves scalar multiplication, matrix inversion, matrix multiplications with values in $H^{s-1}(M)$ wich are smooth operations since $H^{s-1}(M)$ is a Hilbert algebra for $s > d/2 + 1$. Thus, the metric is smooth.

We refer to [49, Theorem 4.1] for a proof of the fact that the exponential map is smooth when $s > d/2 + 2$. □

Consequently, the geodesic equation can be interpreted as an ODE in $H^s(M, M)$ which proves local well-posedness of the geodesic equation. However, geodesic completeness (global well-posedness) does not hold since there exists smooth initial conditions for the Camassa-Holm equation such that the solutions blow up in finite time. As a consequence, metric completeness does not hold either (since it would imply geodesic completeness). The Gauss lemma is valid in this strong H^s topology which ensures in particular that geodesics are length minimizing among all curves that stay in a H^s neighborhood. However, this is *not* enough to prove that the associated geodesic distance is non degenerate since an almost minimizing geodesic can escape this neighborhood for arbitrarily small energy. This is what happens for the right-invariant metric $H^{1/2}$ on the circle S_1 where the metric is degenerate although there exists a smooth exponential map similarly to our case, see [20].

In [47], Michor and Mumford proved that the right-invariant metric L^2 on the group of diffeomorphisms leads to a degenerate distance, i.e. between any two diffeomorphisms, the infimum of the path lengths joining them is zero. This is not the case for the H^{div} right-invariant metric, the following theorem was also proven in their article.

Theorem 28 (Michor and Mumford)**.** *The distance on* $\mathrm{Diff}(M)$ *induced by the* H^{div} *right-invariant metric is non-degenerate. Namely, between two distinct diffeomorphisms the infimum of the lengths of the paths joining them is strictly positive.*

5. A Riemannian submanifold point of view on the H^{div} right-invariant metric

The Riemannian submersion $\pi_0 : \mathrm{Aut}(\mathcal{C}(M)) \mapsto \mathrm{Dens}(M)$ defined in Proposition 11 enables to study the equivalent problem to the incompressible Euler equation. The fiber of the Riemannian submersion at vol is $\pi_0^{-1}(\{\mathrm{vol}\})$ and it will be denoted as in Section 3.2 by $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$. More explicitely, we have

$$\pi_0^{-1}(\{\mathrm{vol}\}) = \{(\varphi, \lambda) \in \mathrm{Aut}(\mathcal{C}(M)) \,:\, \varphi_*(\lambda^2 \,\mathrm{vol}) = \mathrm{vol}\}\,. \tag{5.1}$$

The constraint $\varphi_*(\lambda^2 \,\mathrm{vol}) = \mathrm{vol}$ can be made explicit as follows

$$\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M)) = \{(\varphi, \sqrt{\operatorname{Jac}(\varphi)}) \in \mathrm{Aut}(\mathcal{C}(M)) \,:\, \varphi \in \mathrm{Diff}(M)\}\,. \tag{5.2}$$

Note that this isotropy subgroup can be identified with the group of diffeomorphims of M since the map $\varphi \mapsto (\varphi, \sqrt{\operatorname{Jac}(\varphi)})$ is also a section of the short exact sequence (2.15). This shows that there is a natural identification between $\mathrm{Diff}(M)$ and $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$. Now, the vertical space at point $(\varphi, \sqrt{\operatorname{Jac}(\varphi)}) \in \mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ is

$$\operatorname{Ker}\left(d\pi_0(\varphi, \sqrt{\operatorname{Jac}(\varphi)})\right) = \{(v, \alpha) \cdot (\varphi, \sqrt{\operatorname{Jac}(\varphi)}) \,:\, \operatorname{div} v = 2\alpha\,\}\,, \tag{5.3}$$

and equivalently

$$\operatorname{Ker}\left(d\pi_0(\varphi, \sqrt{\operatorname{Jac}(\varphi)})\right) = \left\{ \left(v, \frac{1}{2}\operatorname{div} v\right) \cdot (\varphi, \sqrt{\operatorname{Jac}(\varphi)}) \,:\, v \in \operatorname{Vect}(M) \right\} . \tag{5.4}$$

The metric $L^2(M, \mathcal{C}(M))$ on $\operatorname{Aut}(\mathcal{C}(M))$ restricted to $\operatorname{Diff}(M) \simeq \operatorname{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ reads

$$G_\varphi(X_\varphi, X_\varphi) = \int_M |v|^2 \,\mathrm{dvol} + \frac{1}{4}\int_M |\operatorname{div} v|^2 \,\mathrm{dvol}\,, \tag{5.5}$$

where $v = X_\varphi \circ \varphi^{-1}$. Therefore, on $\operatorname{Diff}(M) \simeq \operatorname{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$, the induced metric is a right-invariant H^{div} metric. In other words, we have

Theorem 29. *By its identification with* $\operatorname{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$*, the diffeomorphism group endowed with the* H^{div} *right-invariant metric* (2.2) *is isometrically embedded in* $L^2(M, \mathcal{C}(M))$.

As a straightforward application, we retrieve theorem 28.

Corollary 30. *The distance on* $\operatorname{Diff}(M)$ *with the right-invariant metric* H^{div} *is non degenerate.*

Proof. Let $\varphi_0, \varphi_1 \in \operatorname{Diff}(M)$ be two diffeomorphisms and c be a path joining them. The length of the path c for the right-invariant metric H^{div} is equal to the length of the lifted path $\tilde{c}$ in $\operatorname{Aut}(\mathcal{C}(M))$. Since $L^2(M, \mathcal{C}(M))$ is a Hilbert manifold, the length of the path $\tilde{c}$ is bounded below by the length of the geodesic joining the natural lifts of φ_0 and φ_1 in $L^2(M, \mathcal{C}(M))$. Therefore, it leads to

$$d_{H^{\mathrm{div}}}(\varphi_0, \varphi_1) \geq d_{L^2(M,\mathcal{C}(M))}\left((\varphi_0, \sqrt{\operatorname{Jac}(\varphi_0)}), (\varphi_1, \sqrt{\operatorname{Jac}(\varphi_1)})\right) . \tag{5.6}$$

If $d_{H^{\mathrm{div}}}(\varphi_0, \varphi_1) = 0$ then $d_{L^2(M,\mathcal{C}(M))}\left((\varphi_0, \sqrt{\operatorname{Jac}(\varphi_0)}), (\varphi_1, \sqrt{\operatorname{Jac}(\varphi_1)})\right) = 0$ which implies $\varphi_0 = \varphi_1$. □

Remark 6 (The Fisher-Rao metric)**.** *In* [31]*, it is shown that the* $\dot{H}^1$ *right-invariant metric descends to the Fisher-Rao metric on the space of densities. Let us explain why this situation differs from our: It is well known that a left action of a group endowed with a right-invariant metric induces on the orbit a Riemannian metric for which the action is a Riemannian submersion. However, Khesin et al. do not consider a left action, but a right action on the space of densities: More precisely, if a reference density* ρ *is chosen, the map they considered is*

$$\operatorname{Diff}(M) \to \operatorname{Dens}(M)$$
$$\varphi \mapsto \varphi^* \rho\,.$$

Obviously, this situation is equivalent to a left action of a group of diffeomorphisms endowed with a left invariant metric. In such a situation, the descending metric property has to be checked [31, Proposition 2.3].

Their result can be read from our point of view: The $\dot{H}^1$ *metric is* $\frac{1}{4}\int_M |\operatorname{div} v|^2 \,\mathrm{d}\mu$ *and it corresponds to the case where* $a = 0$*. It thus leads to a degenerate metric on the group. Viewed in the ambient space* $L^2(M, \mathcal{C}(M))$*, the projection on the bundle component is a (pseudo-) isometry from* $L^2(M, \mathcal{C}(M))$ *(endowed with this pseudo-metric) to the space of densities since* $a = 0$*. Moreover, on the space of densities which lie in the image of the projection, that is, the set of probability densities, the projected metric is the Fisher-Rao metric.*

We now use the identification between $\operatorname{Diff}(M)$ endowed with the right-invariant H^{div} metric and $\operatorname{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ as a submanifold of $\operatorname{Aut}(\mathcal{C}(M))$ and write the geodesic equations in this setting. As is standard for the incompressible Euler equation, the constraint is written in Eulerian coordinates and the corresponding geodesic are written hereafter.

Theorem 31. *The geodesic equations on the fiber* $\operatorname{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ *as a Riemannian submanifold of* $\operatorname{Aut}(\mathcal{C}(M))$ *endowed with the metric* $L^2(M, \mathcal{C}(M))$ *can be written in Lagrangian coordinates*

$$\begin{cases} \frac{D}{Dt}\dot{\varphi} + 2\frac{\dot{\lambda}}{\lambda}\dot{\varphi} = -\nabla^g P \circ \varphi \\ \ddot{\lambda} - \lambda g(\dot{\varphi}, \dot{\varphi}) = -2\lambda P \circ \varphi\,, \end{cases} \tag{5.7}$$

with a function $P : M \to \mathbb{R}$.
In Eulerian coordinates, the geodesic equations read

$$\begin{cases} \dot{v} + \nabla^g_v v + 2v\alpha = -\nabla^g P \\ \dot{\alpha} + \langle \nabla\alpha, v\rangle + \alpha^2 - g(v,v) = -2P\,, \end{cases} \tag{5.8}$$

where $\alpha = \frac{\dot{\lambda}}{\lambda} \circ \varphi^{-1}$ and $v = \partial_t \varphi \circ \varphi^{-1}$.

This submanifold point of view leads to a generalization of [30, Theorem A.2] on the sectional curvature of $\mathrm{Diff}(M)$ which has been computed and studied in [30]. The authors show that the curvature of $\mathrm{Diff}(S_1)$ can be written using the Gauss-Codazzi formula and they show the explicit embedding in a semi-direct product of groups similar to our situation.

As mentioned above, we consider $\mathrm{Diff}(M)$ as a submanifold of $L^2(M, \mathcal{C}(M))$. The second fundamental form can be computed as in the case of the incompressible Euler equation.

Proposition 32. *Let U, V be two smooth right-invariant vector fields on $\mathrm{Aut}(\mathcal{C}(M))$ that can be written as $U(\varphi,\lambda) = (u,\alpha)\circ(\varphi,\lambda)$ and $V(\varphi,\lambda) = (v,\beta)\circ(\varphi,\lambda)$. The second fundamental form for the isometric embedding $\mathrm{Diff}(M) \hookrightarrow L^2(M,\mathcal{C}(M))$ is*

$$\mathrm{II}(U,V) = (\nabla P \circ \varphi, 2\lambda P \circ \varphi)\ , \tag{5.9}$$

where $P = (2\,\mathrm{Id} - \Delta)^{-1} A(\nabla_{(u,\alpha)}(v,\beta))$ is the unique solution of the elliptic PDE (2.26)

$$(2\,\mathrm{Id} - \Delta)(P) = A(\nabla_{(u,\alpha)}(v,\beta))\,, \tag{5.10}$$

where $A(w,\gamma) \stackrel{\text{def.}}{=} -\operatorname{div}(w) + \gamma$. Using the explicit expression of $\nabla_{(u,\alpha)}(v,\beta)$ the elliptic PDE reads

$$(2\,\mathrm{Id} - \Delta)(P) = -\operatorname{div}(\nabla_u v + \beta u + \alpha v) + 2\langle \nabla\beta, u\rangle - 2g(u,v) + 2\alpha\beta\,. \tag{5.11}$$

Proof. By right-invariance of the metric, it suffices to treat the case $(\varphi,\lambda) = \mathrm{Id}$. The orthogonal projection is the horizontal lift defined in Proposition 12. Therefore, we compute the infinitesimal action of $\nabla_{(u,\alpha)}(v,\beta)$ on the volume form which is given by the linear operator A and we consider its horizontal lift $(\nabla P, 2P)$ given by Proposition 12. Then, the orthogonal part of $\nabla_{(u,\alpha)}(v,\beta)$ to the tangent space of $\mathrm{Diff}(M)$ at Id is given by $(\nabla P, 2\lambda P)$. By right-invariance, the orthogonal projection at (φ,λ) is given by $(\nabla P \circ \varphi, 2\lambda P \circ \varphi)$.

From Proposition 4, one has

$$\nabla_{(u,\alpha)}(v,\beta) = (\nabla_u v + \beta u + \alpha v, \langle \nabla\beta, u\rangle - g(u,v) + \alpha\beta)\ , \tag{5.12}$$

and Formula (5.11) follows. ☐

We can then state the Gauss-Codazzi formula applied to our context.

Proposition 33. *Let U, V be two smooth right-invariant vector fields on $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ written as $U(\varphi,\lambda) = (u,\alpha)\circ(\varphi,\lambda)$ and $V(\varphi,\lambda) = (v,\beta)\circ(\varphi,\lambda)$. The sectional curvature of $\mathrm{Diff}(M)$ endowed with the right-invariant H^{div} metric is*

$$\langle R_{\mathrm{Diff}(M)}(U,V)V,U\rangle = \langle R_{L^2(M,\mathcal{C}(M))}(U,V)V,U\rangle + \langle \mathrm{II}(U,U), \mathrm{II}(V,V)\rangle - \langle \mathrm{II}(U,V), \mathrm{II}(U,V)\rangle\,. \tag{5.13}$$

where II is the second fundamental form (5.9) and

$$\langle R_{L^2(M,\mathcal{C}(M))}(U,V)V,U\rangle = \int_M \langle R_{\mathcal{C}(M)}(u,v)v,u\rangle \circ (\varphi,\lambda)\,\mathrm{d}\mu\,, \tag{5.14}$$

where $(\varphi,\lambda) \in \mathrm{Aut}(\mathcal{C}(M))$.

Proof. The only remaining point is the computation of the sectional curvature of $L^2(M,\mathcal{C}(M))$ which is done in Freed and Groisser's article [22]. ☐

Note that the the sectional curvature of $L^2(M,\mathcal{C}(M))$ vanishes if $M = S_n$ since $\mathcal{C}(M) = \mathbb{R}^{n+1}$, which is the case for the one-dimensional Camassa-Holm equation. However, for $M = T_n$, $n \geq 2$, the flat torus, the sectional curvature of $\mathcal{C}(M)$ is non-positive and bounded below by -1 and thus the sectional curvature of $L^2(T_n,\mathcal{C}(T_n))$ is non-positive.

6. Applications

The point of view developed above provides an example of an isometric embedding of the group of diffeomorphisms endowed with the right-invariant H^{div} metric in an L^2 space such as $L^2(M,N)$, here with $N = \mathcal{C}(M)$. This may bring additional informations to the understanding of the corresponding fluid dynamic equation. Let us detail the one-dimensional situation, that is when $M = S_1(r)$, the circle of radius r (we denote simply denote it S_1 if $r = 1$). In such a case, $\mathcal{C}(S_1(r))$ is locally flat and all the (singular) curvature concentrates at the cone point. Note that, when $r < 1$ (or equivalently, its Riemannian diameter less than π), the cone is a usual cone embedded in $\mathbb{R}^3$. Last, $\mathcal{C}(S_1)$ is isometric to $\mathbb{R}^2 \setminus \{0\}$, actually one can add the cone point which turns the isometry into an isometry between metric spaces. We use such an isometry to define, following Theorem 29,

$$\mathcal{M} : \mathrm{Diff}(S_1) \to \mathrm{Aut}(\mathcal{C}(S_1)) \subset L^2(S_1, \mathbb{R}^2) \tag{6.1}$$

$$\varphi \mapsto (\varphi, \sqrt{\varphi'}) = \sqrt{\varphi'} e^{i\varphi} \,. \tag{6.2}$$

Then, the solutions of the Camassa-Holm are geodesics on the isotropy subgroup explicitly written in (5.2). Note that the map $\mathcal{M}$ is very similar to a Madelung transform which maps solutions of the Schrödinger equation to solutions of a compressible Euler type of hydrodynamical equation. In our case, we show that $\mathcal{M}$ maps solutions of the Camassa-Holm equation to solutions of the incompressible Euler equation on the plane $\mathbb{R}^2 \setminus \{0\}$ for a density which has a singularity at 0. Note that this discussion generalizes directly to the case $M = S_n$ since $\mathcal{C}(M)$ is isometric to $\mathbb{R}^{n+1}$. In the general case, we are left with the geometry of the cone, and therefore, the map $\mathcal{M}$ maps solutions of the geodesic equation on the diffeomorphisms group for the right-invariant H^{div} metric to solutions of the incompressible Euler equation on the $\mathcal{C}(M)$ for a density which has a singularity at the cone point.

In this section, we present this result in a general setting and we apply this Riemannian submanifold point of view to derive a similar result to Brenier, namely that smooth geodesics are length minimizing for short times.

6.1. The Camassa-Holm equation as an Euler equation on the cone.

Formula (6.4) is close to the incompressible Euler equation in Lagrangian coordinates. However, the geodesic equation (6.4) is apparently written on the space of maps $M \mapsto \mathcal{C}(M)$. Since $\mathrm{Aut}(\mathcal{C}(M)) \subset \mathrm{Diff}(\mathcal{C}(M))$, this geodesic equation can be expected to be a geodesic equation on the group of diffeomorphism of the cone. This is indeed the case, the second equation in (5.7) being linear in λ and the first equation being 0 homogeneous in λ, the geodesic equation can be rewritten as

$$\begin{cases} \frac{D}{Dt}\dot\varphi + 2\frac{\dot\lambda}{\lambda}\dot\varphi = -\nabla^g P \circ \varphi \\ \ddot\lambda r - \lambda r g(\dot\varphi, \dot\varphi) = -2\lambda r P \circ \varphi \,. \end{cases} \tag{6.3}$$

However, the diffeomorphisms $(\varphi, \lambda) \in \mathrm{Diff}(\mathcal{C}(M))$ do not preserve the Riemannian volume measure on $\mathcal{C}(M)$ but another density which has a singularity at the cone point. This amounts to rewriting the left action defined by π in (2.3) as a pushforward of a density on the cone.

Theorem 34. *On the group of diffeomorphisms of the cone, the geodesic equation can be written*

$$\frac{D}{Dt}(\dot\varphi, \dot\lambda r) = -\nabla \Psi_P \circ (\varphi, \lambda r) \,, \tag{6.4}$$

where $\Psi_P(x,r) \overset{\text{def.}}{=} r^2 P(x)$. *Moreover, the diffeomorphisms of* $\mathcal{C}(M)$ (φ,λ) *preserve the measure* $\tilde\nu \overset{\text{def.}}{=} r^{-3}\,\mathrm{d}r\,\mathrm{dvol}$.
In other words, a solution (φ,λ) *of* (6.4) *is a solution of the incompressible Euler equation for the density* $r^{-3-d}\,\mathrm{dvol}_{\mathcal{C}(M)}$ *where* $\mathrm{dvol}_{\mathcal{C}(M)}$ *is the volume form on the cone* $\mathcal{C}(M)$ *and* d *is the dimension of* M.

Proof. The geodesic equations (6.3) can be rewritten in the form (6.4) since a direct computation gives $\nabla \Psi_P = (\nabla^g P, 2rP)$.

The only remaining point is that (φ, λ) preserves the measure $r^{-3}\,\mathrm{d}\nu\,\mathrm{d}r$ on $\mathcal{C}(M)$, if the relation $\lambda = \sqrt{\mathrm{Jac}(\varphi)}$ holds. Indeed, the volume form $r^\alpha\,\mathrm{d}\nu\,\mathrm{d}r$ is preserved by (φ, λ) if and only if the following equality is satisfied $(\lambda r)^\alpha \lambda\,\mathrm{Jac}(\varphi) = r^\alpha$, equivalently $\lambda^{\alpha+3} = 1$. It is the case if and only if $\alpha = -3$. $\square$

In particular, this theorem underlines that $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M)) = \mathrm{Aut}(\mathcal{C}(M)) \cap \mathrm{SDiff}_{\tilde{\nu}}(\mathcal{C}(M))$. In remark 2, we mentioned that $\mathrm{Aut}(\mathcal{C}(M))$ is a totally geodesic subspace of $\mathrm{Diff}(\mathcal{C}(M))$, which explains the fact that the geodesic equation on $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ is actually a geodesic equation on $\mathrm{SDiff}_{\tilde{\nu}}(\mathcal{C}(M))$. We illustrate this situation in Figure 1.

Let us be interested in the particular case when $M = S_1$. The Camassa-Holm equation is

$$a^2 \partial_t u - b^2 \partial_{txx} u + 3a^2 \partial_x u\, u - 2b^2 \partial_{xx} u\, \partial_x u - b^2 \partial_{xxx} u\, u = 0\,. \tag{6.5}$$

We consider the case $a = 1$ and $b = \frac{1}{2}$ since it gives back the standard metric on $\mathcal{C}(S_1)$. In this case, one has

$$\begin{cases} \partial_t u - \frac{1}{4}\partial_{txx} u + 3\partial_x u\, u - \frac{1}{2}\partial_{xx} u\, \partial_x u - \frac{1}{4}\partial_{xxx} u\, u = 0 \\ \partial_t \varphi(t,x) = u(t, \varphi(t,x))\,. \end{cases} \tag{6.6}$$

Corollary 35. *The solutions of the Camassa-Holm equation (6.6) are mapped by*

$$\mathcal{M}(\varphi) = \left[(r,\theta) \mapsto r\sqrt{\varphi'(\theta)}e^{i\varphi(\theta)}\right] \tag{6.7}$$

to solutions of the incompressible Euler equation for the measure $\tilde{\nu} \overset{\text{def.}}{=} (1/r^4)\,\mathrm{Leb}$ *on* $\mathbb{R}^2 \setminus \{0\}$. *More precisely,* $\mathcal{M}(\varphi)$ *is a solution to equation (1.10) and* $\mathcal{M}(\varphi)$ *lies in the group of* $\tilde{\nu}$*-preserving diffeomorphisms, denoted by* $\mathrm{Diff}_{\tilde{\nu}}(\mathbb{R}^2 \setminus \{0\})$.

The diffeomorphism $\mathcal{M}(\varphi)$ can be extended continuously to $\mathbb{R}^2$, it has to fix the cone point 0. However, in general, it is not a diffeomorphism any longer and only a homeomorphism.

6.2. **Length minimizing geodesics on H^{div}.** We now show that every smooth geodesics are length minimizing on a sufficiently short time interval. This is actually a straightforward generalization of Brenier's proof in the case of Euler equation to a Riemannian setting. Note that the Ebin and Marsden's point of view does not give such precise results since it requires a strong ambient topology for the Gauss lemma to apply; for the H^{div} right-invariant metric, geodesics are length minimizing among paths that lie in an H^s neighborhood for $s > d/2 + 2$. In the worst case of our theorem, we require only an L^∞ bound on the Jacobian and on the diffeomorphism.

Theorem 36. *Let* $(\varphi(t), r(t))$ *be a smooth solution to the geodesic equations (6.4) on the time interval* $[t_0, t_1]$. *If* $(t_1 - t_0)^2 \langle w, \nabla^2 \Psi_{P(t)}(x,r) w\rangle < \pi^2 \|w\|^2$ *holds for all* $t \in [t_0, t_1]$ *and* $(x, r) \in \mathcal{C}(M)$ *and* $w \in T_{(x,r)}\mathcal{C}(M)$, *then for every smooth curve* $(\varphi_0(t), r_0(t)) \in \mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ *satisfying* $(\varphi_0(t_i), r_0(t_i)) = (\varphi(t_i), r(t_i))$ *for* $i = 0, 1$ *and the condition* $(*)$, *one has*

$$\int_{t_0}^{t_1} \|(\dot\varphi, \dot r)\|^2 \,\mathrm{d}t \le \int_{t_0}^{t_1} \|(\dot\varphi_0, \dot r_0)\|^2 \,\mathrm{d}t\,, \tag{6.8}$$

with equality if and only if the two paths coincide on $[t_0, t_1]$.

Define $\delta_0 \overset{\text{def.}}{=} \min\{r(x,t) \,:\, \text{injectivity radius at } (\varphi(t,x), r(t,x))\}$, *then the condition* $(*)$ *is:*

1. *If the sectional curvature of* $\mathcal{C}(M)$ *can assume both signs or if* $\mathrm{diam}(M) \ge \pi$, *there exists* δ *satisfying* $0 < \delta < \delta_0$ *such that the curve* $(\varphi_0(t), r_0(t))$ *has to belong to a* δ*-neighborhood of* $(\varphi(t), r(t))$, *namely*
$$d_{\mathcal{C}(M)}\left((\varphi_0(t,x), r_0(t,x)), (\varphi(t,x), r(t,x))\right) \le \delta$$
for all $(x,t) \in M \times [t_0, t_1]$ *where* $d_{\mathcal{C}(M)}$ *is the distance on the cone.*
2. *If* $\mathcal{C}(M)$ *has non positive sectional curvature, then, for every* $\delta < \delta_0$, *there exists a short enough time interval on which the geodesic will be length minimizing.*
3. *If* $M = S_d(1)$, *the result is valid for every path* $(\dot\varphi_0, \dot r_0)$.

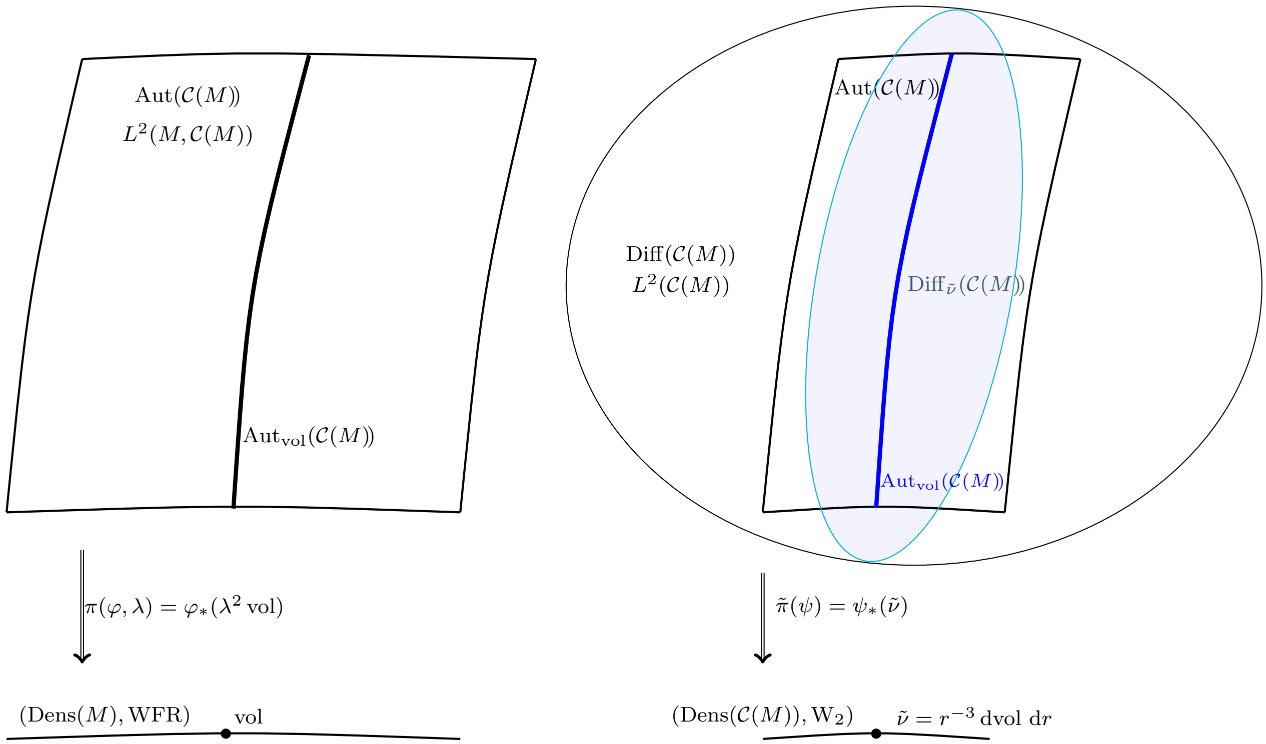

FIGURE 1. On the left, the picture represents the Riemannian submersion between $\mathrm{Aut}(\mathcal{C}(M))$ and the space of positive densities on M and the fiber above the volume form is $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$. On the right, the picture represents the automorphism group $\mathrm{Aut}(\mathcal{C}(M))$ isometrically embedded in $\mathrm{Diff}(\mathcal{C}(M))$ and the intersection of $\mathrm{Diff}_{\tilde{\nu}}(\mathcal{C}(M))$ and $\mathrm{Aut}(\mathcal{C}(M))$ is equal to $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$.

Importantly, the condition on the Hessian is not empty, i.e. it is fulfilled in our case of interest: Indeed, when P is a C^2 function on M, the Hessian of $\Psi_P(x,r) = r^2 P(x)$ is, in the orthonormal basis $\partial_r, \frac{1}{r}e_1, \ldots, \frac{1}{r}e_d$ where $e_1, \ldots, e_d$ is an orthornormal basis of T_xM

$$\nabla^2 \Psi_P(x,r) = \begin{pmatrix} \nabla^2 P(x) & 2\nabla P(x) \\ 2\nabla P^T(x) & 2P(x) \end{pmatrix}, \tag{6.9}$$

where ∇P is the gradient of P in the orthonormal basis $e_1, \ldots, e_d$. Since P is smooth and M is compact, the Hessian of P is bounded uniformly on $\mathcal{C}(M)$.

Proof. To alleviate notations, we denote $g_t = (\varphi(t), r(t))$ and $h_t = (\varphi_0(t), r_0(t))$. Since $g_t = (\varphi(t), \sqrt{\mathrm{Jac}(\varphi(t))})$, by direct integration, for every $t \in [t_0, t_1]$

$$\int_M \Psi_P(g_t(s))\,\mathrm{d}s = 0\,, \tag{6.10}$$

and the same equality holds for h_t.

Let $s \in [0,1] \mapsto c(t,s,x)$ be a two parameters ($t \in [t_0, t_1]$ and $x \in M$) family of geodesics on $\mathcal{C}(M)$ such that $c(t,0,x) = g_t(x)$ and $c(t,1,x) = h_t(x)$ for every $t \in [t_0, t_1]$ and $x \in M$. This family of geodesics is uniquely defined if one considers balls which do not intersect the cut locus. Uniformity of the radius of the balls can be obtained since $[t_0, t_1] \times M$ is compact, which defines δ_0. Consequently, the family of curves $c(t,s,x)$ is a smooth family of geodesics, at least as smooth as $g_t(x)$ and $h_t(x)$ are with respect to the parameters t, x. Since $\partial_t c(t,s,x)$ is a variation of geodesics, it is a Jacobi field as a function of s. Thus, we will use the notation $J(t,s,x) = \partial_t c(t,s,x)$. Consequently, we have

$$J(t,0,x) = \partial_t g_t(x) \text{ and } J(t,1,x) = \partial_t h_t(x)\,. \tag{6.11}$$

Now, the result we want to prove can be reformulated as,

$$\int_{t_0}^{t_1}\int_M \|J(t,0,x)\|^2\,\mathrm{d}t\,\mathrm{d}x \leq \int_{t_0}^{t_1}\int_M \|J(t,1,x)\|^2\,\mathrm{d}t\,\mathrm{d}x \tag{6.12}$$

with equality if and only if for almost every x, it holds $g_t(x) = h_t(x)$ for all $t \in [t_1,t_2]$. We now use a second-order Taylor expansion of $\Psi_P(c(t,s,x))$ with respect to s at $s=0$. Denoting by $M \overset{\text{def.}}{=} \sup_{t\in[t_0,t_1]}\sup_{x\in M}|\nabla^2\Psi_{P_t}(x)|$, we have, writing $c(s)$ for $c(t,s,x)$,

$$\Psi_P(h_t(x)) - \Psi_P(g_t(x)) - \langle \nabla\Psi_P(c(0)), \partial_s c(0)\rangle \leq \frac{M}{2}\int_0^1 \|\partial_s c(s)\|^2\,\mathrm{d}s\,.$$

Now, one has that $\partial_s c(t,s,x)$ vanishes at $t=0$ and $t=1$. We can therefore apply Poincaré inequality to $\|\partial_s c(s)\|$ to obtain

$$\int_{t_0}^{t_1}\|\partial_s c(s)\|^2\,\mathrm{d}s \leq \frac{M(t_1-t_0)^2}{2\pi^2}\int_{t_0}^{t_1}|\partial_t\|\partial_s c(s)\||^2\,\mathrm{d}s\,. \tag{6.13}$$

Since $\partial_t\|\partial_s c(s)\| = \frac{1}{\|\partial_s c\|}\langle\nabla_t\partial_s c, \partial_s c\rangle$, we have the inequality $|\partial_t\|\partial_s c(s)\|| \leq \|\nabla_t\partial_s c\|$ and we get, exchanging derivatives,

$$\int_{t_0}^{t_1}\|\partial_s c(s)\|^2\,\mathrm{d}s \leq \frac{M(t_1-t_0)^2}{2\pi^2}\int_{t_0}^{t_1}\|\dot{J}(s)\|^2\,\mathrm{d}s\,, \tag{6.14}$$

where $\dot{J}$ is the covariant derivative of J with respect to s. We thus have

$$\int_{t_0}^{t_1}\Psi_P(c(t,1,x)) - \Psi_P(c(t,0,x)) - \langle\nabla\Psi_P(c(t,0,x)), \partial_s c(0)\rangle \leq \frac{M(t_1-t_0)^2}{2\pi^2}\int_{t_0}^{t_1}\|\dot{J}(s)\|^2\,\mathrm{d}s\,.$$

However, $g_t(x) = c(t,0,x)$ is a solution of $\nabla_t\partial_t c(t,0,x) = -\nabla\Psi_P(t,0,x)$, therefore, an integration by part w.r.t. t leads to

$$\int_{t_0}^{t_1}\Psi_P(c(t,1,x)) - \Psi_P(c(t,0,x)) - \langle\partial_t c(t,0,x), \nabla_t\partial_s c(0)\rangle\,\mathrm{d}t \leq \frac{M(t_1-t_0)^2}{2\pi^2}\int_{t_0}^{t_1}\|\dot{J}(s)\|^2\,\mathrm{d}s\,.$$

Last, integrating over M and exchanging once again covariant derivatives gives

$$\int_{t_0}^{t_1}\int_M -\langle J(t,0,x), \dot{J}(t,0,x)\rangle\,\mathrm{d}x\,\mathrm{d}t \leq \frac{M(t_1-t_0)^2}{2\pi^2}\int_{t_0}^{t_1}\int_M\int_0^1\|\dot{J}(t,s,x)\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,.$$

Writing $f(s) = \int_{t_0}^{t_1}\int_M \|J(t,s,x)\|^2\,\mathrm{d}t$, we want to prove $f(1)\geq f(0)$ and we have

$$-f'(0) \leq \frac{M(t_1-t_0)^2}{2\pi^2}\int_{t_0}^{t_1}\int_M\int_0^1\|\dot{J}(t,s,x)\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,.$$

Therefore, the result is proven if we can show

$$f(1) - f(0) - f'(0) \geq \varepsilon\int_{t_0}^{t_1}\int_M\int_0^1\|\dot{J}(t,s,x)\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,. \tag{6.15}$$

The left hand side can be reformulated using $f(1)-f(0)-f'(0) = \int_0^1(1-s)f''(s)\,\mathrm{d}s$ as

$$\int_{t_0}^{t_1}\int_M\int_0^1(1-s)(\|\dot{J}\|^2 - \langle R(\partial_s c, J)J, \partial_s c\rangle)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t \geq \varepsilon\int_{t_0}^{t_1}\int_M\int_0^1\|\dot{J}\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,, \tag{6.16}$$

with $\varepsilon = \frac{M(t_1-t_0)^2}{2\pi^2}$.

We now need to distinguish between two cases, the first one being when $\int_{t_0}^{t_1}\int_M\int_0^1\|\dot{J}\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t \geq 1$. In this case, we use the inequality

$$\|J(t)\|^2 \leq 2\|J(0)\|^2 + 2\int_0^1\|\dot{J}(s)\|^2\,\mathrm{d}s\,, \tag{6.17}$$

in order to get

$$-\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\langle R(\partial_s c,J)J,\partial_s c\rangle\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\leq\delta^2\int_{t_0}^{t_1}\int_M\int_0^1 K_{\sup}(2\|J(0)\|^2+2\|\dot J(s)\|^2)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,, \tag{6.18}$$

where $\delta=\sup_{(x,t)\in M\times[t_0,t_1]}\|\partial_s c(t,0,x)\|$ and $K_{\sup}$ is a bound on $\max(K(y),0)$ with $K(y)$ is the maximum of the sectional curvatures at $y\in\mathcal{C}(M)$ for y in a bounded neighborhood of $\bigcup_{t\in[t_0,t_1]} g_t(M)$ which is compact. Then, there exists δ sufficiently small such that for every $(x,t)\in M\times[t_0,t_1]$,

$$\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\langle R(\partial_s c,J)J,\partial_s c\rangle\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\leq 1\leq\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,. \tag{6.19}$$

Now we study the second case, that is when $\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\leq 1$. Applying once again inequality (6.14), we obtain, using the Cauchy-Schwarz inequality,

$$\begin{aligned}\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\langle R(\partial_s c,J)J,\partial_s c\rangle\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t&\leq\varepsilon K_{\sup}\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\|J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\\&\leq\varepsilon K_{\sup}\left(\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^4\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}\left(\int_{t_0}^{t_1}\int_M\int_0^1\|J\|^4\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right)^{1/2}.\end{aligned} \tag{6.20}$$

We now remark that for each t,x, the space of Jacobi fields is finite dimensional and consequently, norms are equivalent so that there exists a positive constant m that depends on t,x such that

$$\left(\int_0^1\|\dot J\|^4\,\mathrm{d}s\right)^{1/2}\leq m\int_0^1\|\dot J\|^2\,\mathrm{d}s \tag{6.21}$$

and

$$\left(\int_0^1\|J\|^4\,\mathrm{d}s\right)^{1/2}\leq m\int_0^1\|J\|^2\,\mathrm{d}s\,. \tag{6.22}$$

By compactness of $M\times[t_0,t_1]$, the constant m can be chosen independently of t,x and therefore, there exists a constant m' such that

$$\begin{aligned}\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\langle R(\partial_s c,J)J,\partial_s c\rangle\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t&\leq\\ \varepsilon K_{\sup}m'\left(\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right)&\left(\int_{t_0}^{t_1}\int_M\int_0^1\|J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right).\end{aligned} \tag{6.23}$$

Then, inequality (6.17) leads to

$$\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\langle R(\partial_s c,J)J,\partial_s c\rangle\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\;\leq\;\varepsilon K_{\sup}Mm'\left(\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right), \tag{6.24}$$

with $M=\left(\int_{t_0}^{t_1}\int_M 2\|J(0)\|^2+2\int_0^1\|\dot J(s)\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\right)$.

Let us recall that our goal is to prove the existence of $\varepsilon>0$ such that

$$\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\geq\varepsilon\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2+(1-s)\langle R(\partial_s c,J)J,\partial_s c\rangle\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,, \tag{6.25}$$

which, in the first case, reads

$$\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\geq 2\varepsilon\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,, \tag{6.26}$$

and in the second case

$$\int_{t_0}^{t_1}\int_M\int_0^1(1-s)\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\geq\varepsilon(1+K_{\sup}Mm')\int_{t_0}^{t_1}\int_M\int_0^1\|\dot J\|^2\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,. \tag{6.27}$$

The existence of ε follows from the fact that the space of Jacobi fields is finite dimensional and the fact $M \times [t_0, t_1]$ is compact. It thus proves the result in the general case.

When the cone $\mathcal{C}(M)$ has non-positive sectional curvature, $K_{\sup} = 0$ therefore, we only have to prove the existence of ε such that

$$\int_{t_0}^{t_1} \int_M \int_0^1 (1-s)\|\dot{J}\|^2 \,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t \geq \varepsilon \int_{t_0}^{t_1} \int_M \int_0^1 \|\dot{J}\|^2 \,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\,, \tag{6.28}$$

which does not require an a priori bound on the neighborhood.

When $M = S_d(1)$, $\mathcal{C}(M)$ is flat and $\delta_0 = \infty$ and Jacobi fields are constant and the constant ε does not depend on the neighborhood and is equal to $1/2$ as in Brenier's proof. □

This generalization of Brenier's proof is not completely satisfactory in positive curvature or, in the case of negative curvature, because of the injectivity radius bound. In the former case, the constructed interpolating paths have to pass through the cone point and therefore these paths $c(t, s, x)$ are not smooth any longer w.r.t. s and thus Jacobi fields are smooth a priori. These two limitations could probably be overcome using a different strategy than a geodesic homotopy between the two diffeomorphisms. Let us insist on the flat case, that contains the Camassa-Holm equation on S_1:

Corollary 37. *Let $M = S_n(1)$ and $0 < a$. Smooth solutions to the Camassa-Holm equation (4.13) with parameters $a, b = a/2$ are length minimizing for short times.*

When M is the $d \geq 2$ dimensional flat torus, the cone $\mathcal{C}(M)$ has non-positive curvature and depending on its diameter, the first or second condition in $(*)$ apply. In general, to the best of our knowledge, the proof presented above is the first generalization to Riemannian manifolds of Brenier's proof and it might be possible to improve on this result, especially to get rid of the boundedness assumption. We actually conjecture that the result holds true without the boundedness assumption.

7. Future directions

In this article, we have presented the link between the Camassa-Holm equation and the new L^2 Wasserstein optimal transport metric between positive Radon measures. On one side, we contributed to the extension of this optimal transport metric to the case of Riemannian manifolds and we derived a corresponding polar factorization theorem. On the other side, we presented an isometric embedding of the group of diffeomorphism group endowed with the right-invariant H^{div} metric in the space $L^2(M, \mathcal{C}(M))$. This isometric embedding enables to rewrite the Camassa-Holm equation, via a Madelung transform, as an incompressible Euler equation on the cone. In other words, the Camassa-Holm equation is a geodesic flow on $\mathrm{Aut}_{\mathrm{vol}}(\mathcal{C}(M))$ for the L^2 metric. As an application, this has also led to a result on the minimizing property of geodesics. Very few papers have been interested with the actual variational problem of minimizing geodesics for the H^{div} metric in the sense of Brenier [6, 7] which can be addressed from the point of view developed in our article. Following Brenier, we will investigate elsewhere the uniqueness of the pressure as in [4]. This isometric embedding and the polar factorization theorem opens the way to design new numerical simulations of variational solutions of the Camassa-Holm equation, in the direction of [25, 44].

Following the point of view developed in this article, we plan to rewrite other fluid dynamic equations as geodesic equations on a submanifold of a space of maps endowed with an L^2 norm. The result may have, as shown for the Camassa-Holm equation, interesting analytical consequences.

Appendix A. Group action and Riemannian submersions

A.1. **Riemannian submersion.** Let (M, g_M) and (N, g_N) be two Riemannian manifolds and $f : M \mapsto N$ a differentiable mapping.

Definition 10. The map f is a Riemannian submersion if f is a submersion and for any $x \in M$, the map $df_x : \mathrm{Ker}(df_x)^{\perp} \mapsto T_{f(x)}N$ is an isometry.

In such a case, $\mathrm{Vert}_x \stackrel{\text{def.}}{=} \mathrm{Ker}(df(x))$ is called the vertical space and $\mathrm{Hor}_x \stackrel{\text{def.}}{=} \mathrm{Ker}(df(x))^\perp$ is called the horizontal space. The horizontal spaces can be used to lift a vector field Y on N onto a vector field $\tilde{Y}$ on M which is horizontal. More precisely, $\tilde{Y}$ is the unique horizontal vector field such that $df_x(\tilde{Y}(x)) = Y(f(x))$. The first immediate property is that Riemannian submersions are length decreasing.

Proposition 38. *Let f be a Riemannian submersion as defined above and $c_0 : [0,1] \mapsto M$ be a smooth curve. It then defines a smooth curve on N by $c_1 \stackrel{\text{def.}}{=} f \circ c_0$. Then,*

$$g_N(\dot{c}_1, \dot{c}_1) = g_M(p_{\mathrm{Hor}}(\dot{c}_0), p_{\mathrm{Hor}}(\dot{c}_0)) \leq g_M(\dot{c}_0, \dot{c}_0)\,, \tag{A.1}$$

where p_{Hor} is the orthogonal projection on the horizontal space.

Another property of Riemannian submersions is the following:

Proposition 39. *Let f be a Riemannian submersion as defined above. Every geodesic $\gamma(t)$ on M which is horizontal at a given time t, i.e. $\gamma'(t) \in \mathrm{Hor}_{\gamma(t)}$, is horizontal for all time and the length of γ is equal to the length of $f(\gamma)$.*

An important property is the computation of the curvature tensor of N that can be done via O'Neill's formula detailed below (see [24]).

Theorem 40 (O'Neill's formula)**.** *Let f be a Riemannian submersion as defined above and X, Y be two orthonormal vector fields on M with horizontal lifts $\tilde{X}$ and $\tilde{Y}$, then*

$$K_N(X,Y) = K_M(\tilde{X}, \tilde{Y}) + \frac{3}{4}\|\operatorname{vert}([\tilde{X}, \tilde{Y}])\|^2_M\,, \tag{A.2}$$

where K denotes the sectional curvature and vert *the orthogonal projection on the vertical space.*

Appendix B. Other proofs

Below is the proof of Theorem 19 which is an adaptation to the Riemannian case of the proof in [12]. In particular, not all the details of the proof are given since they can be found in [12]. Note also that this proof, under minor adaptations, applies to the standard Wasserstein L^2 metric on Riemannian manifolds, see for instance the comments in [56, Remarks 8.3]. A proof of the standard Wasserstein case is given in [1] which uses the Nash isometric embedding theorem. The proof below does not use it and develop a simple regularization argument which is intrinsic on Riemannian manifolds.

Proof of Theorem 19. The fact that the minimum for S is attained follows by application of the direct method of calculus of variations. The set Γ is weakly closed and the functional is weakly continuous and S is lower semicontinuous. In the following, we denote by $\mathcal{S}^2(\rho_1, \rho_2)$ the minimization of the r.h.s. of (3.6).

Since d is a distance on the cone, one can prove that $\mathcal{S}$ is a distance on the space of nonnegative Radon measures which is continuous w.r.t. the weak-* topology, as done in [12, Theorems 2,3].

On the set of measures that are finite sum of Dirac masses, the minimization problem (3.6) can be reduced to a linear optimization problem in finite dimension. Indeed, the optimal semi-couplings can be proven to have support on the product of the support of ρ_1 and ρ_2. Denoting $\rho_1 = \sum_i a_i \delta_{x_i}$ and $\rho_2 = \sum_j b_j \delta_{y_j}$ for x_i, y_j a finite number of points in M, optimal semi-couplings can be written as $\gamma^k = \sum_{i,j} m^k_{i,j} \delta_{x_i, x_j}$ for $k = 1, 2$. Then, one has

$$\begin{aligned} \mathcal{S}^2(\rho_1, \rho_2) &= \sum_{i,j} d^2\left((x_i, m^1_{i,j}), (y_j, m^2_{i,j})\right) \\ &\geq \sum_{i,j} \mathrm{WF}^2(m^1_{i,j}\delta_{x_i}, m^2_{i,j}\delta_{y_j}) \geq \mathrm{WF}^2(\rho_1, \rho_2)\,, \end{aligned}$$

where the first inequality comes the fact that the distance on the cone (with mass coordinates) for a geodesic $(x(t), m(t))$ is given by the evaluation of WF on the path $m(t)\delta_{x(t)}$. The second inequality is given by subadditivity of WF^2. By density of this set of measures and weak-* continuity of WF and $\mathcal{S}$, one has $\mathcal{S} \geq \mathrm{WF}$.

The reverse inequality follows using the convexity of WF^2. By subadditivity of WF^2, one has, for any positive Radon measure ρ_3

$$\mathrm{WF}^2(\rho_1 + \rho_3, \rho_2 + \rho_3) \leq \mathrm{WF}^2(\rho_1, \rho_2)\,. \tag{B.1}$$

Using the triangular inequality and the fact that the WF metric is bounded above (up to a multiplicative constant) by the Hellinger distance, we also have, for $\varepsilon_1 > 0$

$$\mathrm{WF}(\rho_0, \rho_1) \leq \mathrm{WF}(\rho_0 + \varepsilon_1 \mathrm{vol}, \rho_1 + \varepsilon_1 \mathrm{vol}) + 2\,\mathrm{cst}\,\sqrt{\varepsilon_1}\,. \tag{B.2}$$

Let us be more precise on the previous inequality: Consider now a path ρ, m, μ which is a solution to the continuity equation (2.4), then so is the path $\rho + \varepsilon_1 \mathrm{vol}, \mathsf{m}, \mu$ satisfying the boundary conditions $\rho(0) = \rho_0$, $\rho(1) = \rho_1$. Note that $\varepsilon_1 \mathrm{vol}$ is constant in time and space. In addition, it is obvious that

$$\mathcal{J}(\rho + \varepsilon_1 \mathrm{vol}, \mathsf{m}, \mu) \leq \mathcal{J}(\rho, \mathsf{m}, \mu)\,.$$

To prove the final result, it suffices to prove that $\mathcal{S}(\rho_0 + \varepsilon_1 \mathrm{vol}, \rho_1 + \varepsilon_1 \mathrm{vol}) \leq \mathcal{J}(\rho + \varepsilon \mathrm{vol}, \mathsf{m}, \mu) + \varepsilon_0$ for any $\varepsilon_0 > 0$. This will be done via a smoothing argument which is standard in the Euclidean case using convolution but has never been adapted, to the best of our knowledge, to work on Riemannian manifolds (see [56, Remarks 8.3]).

Our goal is to prove that there exists a path of smooth quantities $(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon)$ for which $\mathcal{J}(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon)$ is close to $\mathcal{J}(\rho, \mathsf{m}, \mu)$ and ρ_ε is strictly positive and the time endpoints of the path are close in the weak-* topology. The conclusion can then be obtained by integrating the flow defined by the vector field $(\mathsf{m}_\varepsilon/\rho_\varepsilon, \mu_\varepsilon/\rho_\varepsilon)$. It gives that $\mathcal{S}(\rho_\varepsilon(0), \rho_\varepsilon(1)) \leq \mathcal{J}(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon)$ and the conclusion is similar to the Euclidean case [12, Theorem 5].

By compactness of M, it is sufficient to locally smooth the path on M by iteration of this smoothing. Therefore, we will work on a chart U around a point $x_0 \in M$. By Moser's lemma, it is possible to choose the chart such that the volume form is the Lebesgue measure.

Averaging over perturbations of identity: We construct perturbations (of compact support) of the identity which will be local translations around x_0 and which will play the role of the translations in the standard convolution formula. We consider a ball $B(x_0, r_0)$ and a function u whose support is contained in $B(x_0, r_0)$ and is constant equal to 1 on $B(x_0, r_1)$ for $0 < r_1 < r_0$. For a given vector $v \in \mathbb{R}^d$, we consider the map $\Phi_v(x) = x + u(x)v$ which is a smooth diffeomorphism. We extend Φ to the whole manifold M by defining it as identity outside of U.

Let $k : \mathbb{R}^{d+1} \to \mathbb{R}_+$ be a smooth symmetric function whose support is contained in the unit ball and such that $\int k(y)\,\mathrm{d}y = 1$ and define for $\varepsilon > 0$, $k_\varepsilon(x) = k(x/\varepsilon)/\varepsilon^{d+1}$ whose support is thus contained in the ball of radius ε. We define the mollifier $\mathsf{k}_\varepsilon \star$ acting on $f \in C([0,1] \times U, \mathbb{R})$ by

$$(\mathsf{k}_\varepsilon \star f)(s,x) = \int_{\mathbb{R}} \int_U k_\varepsilon(s,v) f(t+s, \Phi_v^{-1}(x))\,\mathrm{d}v\,\mathrm{d}s\,, \tag{B.3}$$

which is well defined for ε small enough, extending the function outside the time interval $[0,1]$ as a constant. Moreover, for ε sufficiently small, it coincides with the usual convolution on a neighborhood of x_0. By duality, it is well defined on Radon measures and extends trivially to vector valued measures as follows:

$$(\mathsf{k}_\varepsilon \star \rho)(s,x) = \int_{\mathbb{R}} \int_U k_\varepsilon(s,v) (\Phi_v)_*(\rho(t+s))\,\mathrm{d}v\,\mathrm{d}s\,, \tag{B.4}$$

$$(\mathsf{k}_\varepsilon \star m)(s,x) = \int_{\mathbb{R}} \int_U k_\varepsilon(s,v)\,\mathrm{Ad}^*_{\Phi_v^{-1}}(\mathsf{m}(t+s))\,\mathrm{d}v\,\mathrm{d}s\,. \tag{B.5}$$

We consider the path $(\Phi_v)_*(\rho)$ which satisfies the continuity equation for the triple of measures $\Big((\Phi_v)_*(\rho), \mathrm{Ad}^*_{\Phi_v^{-1}}(\mathsf{m}), (\Phi_v)_*(\mu)\Big)$ and average over v to consider

$$(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon) = (\mathsf{k}_\varepsilon \star \rho, \mathsf{k}_\varepsilon \star \mathsf{m}, \mathsf{k}_\varepsilon \star \mu)\,. \tag{B.6}$$

As a convex combination, this path satisfies the continuity equation and the boundary conditions are close in the weak-* topology when ε tends to 0. An important remark is that, for ε small enough, $\mathsf{k}_\varepsilon \star \mathrm{Ad}^*_{\Phi_v^{-1}}(\mathsf{m})$ reduces to the standard convolution on m in a small neighborhood of x_0 since $D\Phi_v = \mathrm{Id}$ in a neighborhood of x_0 since $u \equiv 1$ on $B(x_0, r_1)$.

Use of convexity of $\mathcal{J}$: For notation convenience, we denote by f the integrand of $\mathcal{J}$ and we make the abuse of notation to use ρ, m, μ instead of their corresponding densities w.r.t. ν a dominating measure.

Under the change of variables $y = \Phi_v^{-1}(x)$ (we use one homogeneity hereafter) leads to

$$\mathcal{J}(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon) = \int_{[0,1]\times M} f\left(x, (\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon)\right) \mathrm{d}\nu(x) \leq \\ \int_{\mathbb{R}} \int_U \int_{[0,1]\times M} k_\varepsilon(s,v) f(\Phi_v(y), (\rho(t+s), D\Phi_v(t,y)\mathsf{m}(t+s), \mu(t+s)))\, \mathrm{d}\nu(t,y)\, \mathrm{d}t\, \mathrm{d}s\, \mathrm{d}v \,. \tag{B.7}$$

Moreover, since the metric g on M is smooth and in particular uniformly continuous on M and since $\|D\Phi_v - \mathrm{Id}\,\| \leq \mathrm{cst}\|v\|$ for a constant that only depends on u, we thus have, for any $\varepsilon_2 > 0$, the existence of $\delta > 0$ such that if $\|v\| \leq \delta$ then,

$$|g(x)(w,w) - g(\Phi_v(x))(D\Phi_v(x)w, D\Phi_v(x)w)| \leq \varepsilon_2\, g(x)(w,w)\,, \tag{B.8}$$

for every $w \in T_x M$. Therefore, a direct estimation leads to

$$\left| \int_{\mathbb{R}\times M} k_\varepsilon(s,v) f(\Phi_v(x), (\rho(t+s), \mathsf{m}(t+s), \mu(t+s)))\, \mathrm{d}\nu(t,x) - \int_{[0,1]\times M} f(x, (\rho(t), \mathsf{m}(t), \mu(t)))\, \mathrm{d}\nu(t,x) \right| \\ \leq \varepsilon_2 \mathcal{J}(\rho, \mathsf{m}, \mu)\,, \tag{B.9}$$

and as a consequence the desired result,

$$\mathcal{J}(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon) \leq \mathcal{J}(\rho, \mathsf{m}, \mu) + \varepsilon_2 \mathcal{J}(\rho, \mathsf{m}, \mu)\,. \tag{B.10}$$

Since this averaging reduces to standard convolution in the coordinate chart U in a small neighborhood of x_0, it implies that $(\rho_\varepsilon, \mathsf{m}_\varepsilon, \mu_\varepsilon)$ is smooth in a neighborhood of x_0 and $\rho_\varepsilon \geq \varepsilon_1\,\mathrm{vol}$. By compactness of M, iterating a finite number of times this argument gives the desired path. ☐

Proof of Proposition 22 (Approximate differentiability). The proof is an adaptation of [38, Theorem 6.7] using arguments in [42, 57]. In particular we use the notation of [38]. Let (z_0, z_1) be a generalized optimal potential pair for $\mathrm{WF}^2(\rho_0, \rho_1)$ and γ an optimal coupling [38, Theorem 6.3]. We define the associated densities $\sigma_i = e^{-z_i}$, $i = 0, 1$. Since ρ_0 and ρ_1 are admissible [38, Theorem 6.3,b] implies $\mathrm{Supp}\left(p^1_*(\gamma) = \gamma_0\right) = \mathrm{Supp}(\rho_0)$ and $\mathrm{Supp}\left(p^2_*(\gamma) = \gamma_1\right) = \mathrm{Supp}(\rho_1)$. Therefore, there exist Borel sets $A_i \subset \mathrm{Supp}(\rho_i)$ with $\rho_i(M \setminus A_i) = 0$ such that

$$\sigma_0(x)\sigma_1(y) \geq \cos^2(d_{\pi/2}(x,y)) \qquad \text{in } A_0 \times A_1\,, \tag{B.11}$$

$$\sigma_0(x)\sigma_1(y) = \cos^2(d_{\pi/2}(x,y)) \qquad \gamma - \text{a.e. in } A_0 \times A_1\,. \tag{B.12}$$

To construct the set of approximate differentiability let

$$A_{1,n} = \{y \in M\,;\, \sigma_1(y) \geq 1/n\}$$

and consider, the function

$$s_{0,n} = \sup_{y \in A_{1,n}} \frac{\cos^2(d_{\pi/2}(x,y))}{\sigma_1(y)}\,.$$

By construction, $s_{0,n}$ is bounded, Lipschitz and thus differentiable vol a.e. Still by definition, we have $\sigma_0 \geq s_{0,n}$ and thus the sets $A_{0,n} = \{x \in M\,;\, \sigma_0(x) = s_{0,n}(x)\}$ are increasing. Since (B.12) is valid γ a.e. the set $\bigcap_{n=1}^\infty (X \setminus A_{0,n})$ is ρ_0 negligible. Let

$$A'_{0,n} = \left\{ x \in A_{0,n}\,;\, \lim_{r\to 0} \frac{\mathrm{vol}(B(x,r) \cap A_{0,n})}{\mathrm{vol}(B(x,r))} = 1 \text{ and } s_{0,n} \text{is differentiable at } x \right\}$$

be the set of points of $A_{0,n}$ with vol density 1. Remark that $\bigcap_{n=1}^{\infty}(X \setminus A'_{0,n})$ is also ρ_0 negligible. Let $(\bar{x},\bar{y}) \in A'_{0,n} \times A_{1,n}$ be such that

$$s_{0,n}(\bar{x})\sigma_1(\bar{y}) = \cos^2(d_{\pi/2}(\bar{x},\bar{y})) = \sigma_0(\bar{x})\sigma_1(\bar{y})\,.$$

Using (B.11), it holds, for all $x \in A_1$

$$\sigma_1(y) \geq \cos^2(d_{\pi/2}(x,\bar{y}))/s_{0,n}(x)\,.$$

In particular, $\cos^2(d_{\pi/2}(x,\bar{y}))/s_{0,n}(x)$ achieves its maximum at $\bar{x}$, implying $0 \in \nabla^+_{\bar{x}}(\cos^2(d_{\pi/2}(\cdot,\bar{y}))/s_{0,n}(\cdot))$. Since $s_{0,n}$ is differentiable at $\bar{x}$, it yields that $d^2(\cdot,y)$ is super-differentiable. By Lemma 21, it is also sub-differentiable and thus differentiable at $\bar{x}$. It holds

$$0 = \nabla \cos^2\left(\sqrt{2}\sqrt{\frac{1}{2}d^2_{\pi/2}(\bar{x},\bar{y})}\right)/s_{0,n}(\bar{x}) - \cos^2(d_{\pi/2}(\bar{x},\bar{y}))\nabla s_{0,n}(\bar{x})/s^2_{0,n}(\bar{x}) \tag{B.13}$$

$$= -2\sqrt{2}\tan(d_{\pi/2}(\bar{x},\bar{y}))\frac{\sqrt{2}}{2d_{\pi/2}(\bar{x},\bar{y})}\nabla\left(\frac{1}{2}d^2_{\pi/2}(\bar{x},\bar{y})\right) - \nabla \ln s_{0,n}(\bar{x})\,. \tag{B.14}$$

Let $-\nabla\left(\frac{1}{2}d^2_{\pi/2}(\bar{x},\bar{y})\right) = v_{\bar{x}\to\bar{y}} \in T_{\bar{x}}M$ be the unique vector such that $\bar{y} = \exp^M_{\bar{x}}(v_{\bar{x}\to\bar{y}})$, the last equality reads

$$\tilde{\nabla} z_0(\bar{x}) = -\tilde{\nabla}\ln\sigma_0(\bar{x}) = -\nabla \ln s_{0,n}(\bar{x}) = -2\tan(\|v_{\bar{x}\to\bar{y}}\|)\frac{v_{\bar{x}\to\bar{y}}}{\|v_{\bar{x}\to\bar{y}}\|}\,.$$

Therefore, $\bar{y}$ is unique ρ_1 a.e. and given by

$$\bar{y} = \exp^M_{\bar{x}}(v_{\bar{x}\to\bar{y}}) = \exp^M_{\bar{x}}\left(-\arctan\left(\frac{\|\tilde{\nabla} z_0(\bar{x})\|}{2}\right)\frac{\tilde{\nabla} z_0(\bar{x})}{\|\tilde{\nabla} z_0(\bar{x})\|}\right) = \varphi(\bar{x})\,.$$

It implies that γ is concentrated on the graph of φ in particular $\gamma = (Id,\varphi)_*\gamma_0$ and $\varphi_*\gamma_0 = \gamma_1$. The strict convexity of KL implies that the marginals γ_0 and γ_1 are unique [38, Theorem 6.7] thus

$$z_0 = -\log(\sigma_0) = -\log(\frac{\mathrm{d}\gamma_0}{\mathrm{d}\rho_0})$$

is unique ρ_0 a.e. and γ is also unique. Note that we used the admissible condition to say that σ_0 is ρ_0 a.e. positive. In order to prove (3.14), we start from (3.11) and a direct computation yields

(B.15)

$$\begin{aligned}
\mathrm{WF}^2(\rho_0,\rho_1) &= \mathrm{KL}(\gamma_0,\rho_0) + \mathrm{KL}(\gamma_1,\rho_1) + \int_{M^2} c(x,y)\,\mathrm{d}\gamma(x,y)\\
&= \int_M \log\left(e^{-z_0}\right)e^{-z_0}\,\mathrm{d}\rho_0 + \int_M (1-e^{-z_0})\,\mathrm{d}\rho_0 + \int_M \log\left(e^{-z_1}\right)e^{-z_1}\,\mathrm{d}\rho_1 + \int_M (1-e^{-z_1})\,\mathrm{d}\rho_1\\
&+ \int_{M^2} c(x,\varphi(x))\,\mathrm{d}\gamma(x)\\
&= \int_M (1-e^{-z_0})\,\mathrm{d}\rho_0 + \int_M (1-e^{-z_1})\,\mathrm{d}\rho_1 + \int_M [c(x,\varphi(x)) - z_0(x) - z_1(\varphi(x))]\,\mathrm{d}\gamma_0(x)\\
&= \int_M (1-e^{-z_0})\,\mathrm{d}\rho_0 + \int_M (1-e^{-z_1})\,\mathrm{d}\rho_1.
\end{aligned}$$

□

Acknowledgements

We would like to thank Yann Brenier and Klas Modin for stimulating discussions.

References

[1] L. Ambrosio and N. Gigli. *A user's guide to optimal transport*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2013.

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

[3] J-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[4] Y. Brenier. The dual least action problem for an ideal, incompressible fluid. *Archive for Rational Mechanics and Analysis*, 122(4):323–351, 1993.

[5] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[6] Yann Brenier. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm. Pure Appl. Math.*, 52(4):411–452, 1999.

[7] Yann Brenier. Remarks on the minimizing geodesic problem in inviscid incompressible fluid mechanics. *Calc. Var. Partial Differential Equations*, 47(1-2):55–64, 2013.

[8] Alberto Bressan and Adrian Constantin. Global conservative solutions of the Camassa-Holm equation. *Arch. Ration. Mech. Anal.*, 183(2):215–239, 2007.

[9] Alberto Bressan and Massimo Fonte. An optimal transportation metric for solutions of the Camassa-Holm equation. *Methods Appl. Anal.*, 12(2):191–219, 2005.

[10] D. Burago, Y. Burago, and S. Ivanov. A course in metric geometry. *American Mathematical Soc.*, 2001.

[11] Roberto Camassa and Darryl D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71(11):1661–1664, 1993.

[12] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard. Unbalanced Optimal Transport: Geometry and Kantorovich Formulation. *ArXiv e-prints*, August 2015.

[13] L. Chizat, B. Schmitzer, G. Peyré, and F.-X. Vialard. An Interpolating Distance between Optimal Transport and Fisher-Rao. *Found. Comp. Math.*, 2016.

[14] Adrian Constantin. On the scattering problem for the Camassa-Holm equation. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 457(2008):953–970, 2001.

[15] Adrian Constantin and Joachim Escher. Wave breaking for nonlinear nonlocal shallow water equations. *Acta Math.*, 181(2):229–243, 1998.

[16] Adrian Constantin and David Lannes. The hydrodynamical relevance of the camassa–holm and degasperis–procesi equations. *Archive for Rational Mechanics and Analysis*, 192(1):165–186, 2008.

[17] Raphaël Danchin. A few remarks on the Camassa-Holm equation. *Differential Integral Equations*, 14(8):953–988, 2001.

[18] G. De Philippis and A. Figalli. The Monge-Ampère equation and its link to optimal transportation. *Bull. Amer. Math. Soc.*, 51:527–580, 2014.

[19] David G. Ebin and Jerrold Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)*, 92:102–163, 1970.

[20] J. Escher and B. Kolev. Right-invariant Sobolev metrics of fractional order on the diffeomorphism group of the circle. *Journal of Geometric Mechanics*, 6(3):335 – 372, September 2014.

[21] Joachim Escher and Boris Kolev. The degasperis–procesi equation as a non-metric euler equation. *Mathematische Zeitschrift*, 269(3):1137–1153, 2011.

[22] D. S. Freed and D. Groisser. The basic geometry of the manifold of riemannian metrics and of its quotient by the diffeomorphism group. *Michigan Math. J.*, 36(3):323–344, 1989.

[23] S. Gallot. Équations différentielles caractéristiques de la sphère. *Annales scientifiques de l'Ecole Normale Superieure*, 12(2):235–267, 1979.

[24] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Universitext. Springer, 2004.

[25] T. O. Gallouët and Q. Mérigot. A Lagrangian scheme for the incompressible Euler equation using optimal transport. ArXiv e-prints, May 2016.

[26] T. O. Gallouët and L. Monsaingeon. A JKO splitting scheme for Kantorovich-Fisher-Rao gradient flows, 2016.

[27] F. Gay-Balmaz, C. Tronci, and C. Vizman. Geometric dynamics on the automorphism group of principal bundles: geodesic flows, dual pairs and chromomorphism groups. *Journal of Geometric Mechanics*, 5:39–84, 2013.

[28] Katrin Grunert, Helge Holden, and Xavier Raynaud. Lipschitz metric for the periodic camassa–holm equation. *Journal of Differential Equations*, 250(3):1460 – 1492, 2011.

[29] D. D. Holm, J. E. Marsden, and T. S. Ratiu. The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.*, 137:1–81, 1998.

[30] B. Khesin, J. Lenells, G. Misiolek, and S. C. Preston. Curvatures of Sobolev metrics on diffeomorphism groups. *Pure and Applied Mathematics Quarterly*, 9(2):291 – 332, 2013.

[31] B. Khesin, J. Lenells, G. Misiołek, and S. C. Preston. Geometry of Diffeomorphism Groups, Complete integrability and Geometric statistics. *Geom. Funct. Anal.*, 23(1):334–366, 2013.

[32] B. Khesin and R. Wendt. *The geometry of infinite-dimensional groups*, volume 51. Springer Science & Business Media, 2008.

[33] I. Kolář, P. W. Michor, and J. Slovák. *Natural operations in differential geometry*. Springer-Verlag, Berlin, 1993.
[34] S. Kondratyev, L. Monsaingeon, and D. Vorotnikov. A new optimal trasnport distance on the space of finite Radon measures. *Adv. Differential Equations*, 21(11):1117–1164, 2016.
[35] Shinar Kouranbaeva. The Camassa-Holm equation as a geodesic flow on the diffeomorphism group. *J. Math. Phys.*, 40(2):857–868, 1999.
[36] Serge Lang. *Fundamentals of differential geometry*, volume 191 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999.
[37] Jonatan Lenells. Conservation laws of the Camassa-Holm equation. *J. Phys. A, Math. Gen.*, 38(4):869–880, 2005.
[38] M. Liero, A. Mielke, and G. Savaré. Optimal Entropy-Transport problems and a new Hellinger-Kantorovich distance between positive measures. *ArXiv e-prints*, August 2015.
[39] M. Liero, A. Mielke, and G. Savaré. Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. *SIAM J. Image Analysis*, 48(4):2869–2911, 2016.
[40] J. Lott. Some geometric calculations on Wasserstein space. *Communications in Mathematical Physics*, 277(2):423–437, 2008.
[41] J. Maas, M. Rumpf, C. Schönlieb, and S. Simon. A generalized model for optimal transport of images including dissipation and density modulation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6), Apr 2015. arXiv:1504.01988.
[42] R.J. McCann. Polar factorization of maps on riemannian manifolds. *Geometric & Functional Analysis GAFA*, 11(3):589–608, 2001.
[43] Henry P. McKean. Breakdown of the Camassa-Holm equation. *Comm. Pure Appl. Math.*, 57(3):416–418, 2004.
[44] Q. Mérigot and J.-M. Mirebeau. Minimal geodesics along volume preserving maps, through semi-discrete optimal transport. *ArXiv e-prints*, May 2015.
[45] M. Micheli, P. W. Michor, and D. Mumford. Sobolev metrics on diffeomorphism groups and the derived geometry of spaces of submanifolds. *Izvestiya: Mathematics*, 77(3):541, 2013.
[46] P. W. Michor. *Topics in Differential Geometry*, volume 93 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.
[47] Peter W. Michor and David Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Doc. Math.*, 10:217–245, 2005.
[48] G. Misiolek. Classical solutions of the periodic Camassa-Holm equation. *Geometric & Functional Analysis GAFA*, 12(5):1080–1104, 2002.
[49] Gerard Misiolek and Stephen C. Preston. Fredholm properties of riemannian exponential maps on diffeomorphism groups. *Inventiones mathematicae*, 179(1):191–227, 2010.
[50] K. Modin. Generalised Hunter-Saxton equations, optimal information transport, and factorisation of diffeomorphisms. *Journal of Geometric Analysis*, 25(2):1306–1334, April 2015.
[51] J. Moser. On the volume elements on a manifold. *Trans. Amer. Math. Soc.*, 120:286–294, 1965.
[52] Felix Otto. The geometry of dissipative evolution equations: The porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174, 2001.
[53] F Rezakhanlou. Optimal transport problems for contact structures, 2015.
[54] R.T. Rockafellar. Integrals which are convex functionals. II. *Pacific Journal of Mathematics*, 39(2):439–469, 1971.
[55] A. Trouvé and L. Younes. Metamorphoses through lie group action. *Foundations of Computational Mathematics*, 5(2):173–198, 2005.
[56] C. Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.
[57] C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

CMLS, UMR 7640, École Polytechnique, FR-91128 Palaiseau Cedex.
E-mail address: `thomas.gallouet@polytechnique.edu`

Université Paris-Dauphine, PSL Research University, Ceremade, INRIA, Project team Mokaplan
E-mail address: `fxvialard@normalesup.org`