PROJECTIONS ONTO THE CONE OF OPTIMAL TRANSPORT MAPS AND COMPRESSIBLE FLUID FLOWS

MICHAEL WESTDICKENBERG

ABSTRACT. The system of isentropic Euler equations in the potential flow regime can be considered formally as a second order ordinary differential equation on the Wasserstein space of probability measures. This interpretation can be used to derive a variational time discretization. We prove that the approximate solutions generated by this discretization converge to a measure-valued solution of the isentropic Euler equations. The key ingredient is a characterization of the polar cone to the cone of optimal transport maps.

CONTENTS

1. INTRODUCTION

We are interested in compressible fluid flows in the isentropic regime, where the thermodynamical entropy is assumed to be constant throughout time and space. In this case, the compressible Euler equations take the following form:

$$\left.\begin{aligned} \partial_t \varrho + \nabla\cdot(\varrho\mathbf{u}) &= 0 \\ \partial_t(\varrho\mathbf{u}) + \nabla\cdot(\varrho\mathbf{u}\otimes\mathbf{u}) + \nabla P(\varrho) &= 0 \end{aligned}\right\} \quad \text{in } [0,\infty)\times\mathbf{R}^d, \tag{1.1}$$

for the unknowns $(\varrho, \mathbf{u})\colon [0,\infty)\times\mathbf{R}^d \longrightarrow \mathbb{U}$ with $\mathbb{U} := [0,\infty)\times\mathbf{R}^d$. The nonnegative function ϱ is called the density; it describes the distribution of mass in space and time. The $\mathbf{R}^d$-valued function $\mathbf{u}$ is the Eulerian velocity field. The pressure P is related to the internal energy U of the fluid via the relation

$$P(r) = U'(r)r - U(r) \quad \text{for all } r \geqslant 0. \tag{1.2}$$

We will mostly consider the case of polytropic gases, for which

$$P(r) = \kappa r^\gamma \quad \text{for all } r \geqslant 0, \tag{1.3}$$

Date: January 21, 2012.
2000 *Mathematics Subject Classification.* 35L65, 49J40, 82C40.
Key words and phrases. Optimal Transport, Isentropic Euler Equations.

with constants $\gamma > 1$ and $\kappa \geqslant 0$. We assume that

$$(\varrho, \mathbf{u})(0, \cdot) = (\bar{\varrho}, \bar{\mathbf{u}}) \quad \text{for suitable initial data } (\bar{\varrho}, \bar{\mathbf{u}}) \colon \mathbf{R}^d \longrightarrow \mathbb{U}.$$

For smooth density/velocity $(\varrho, \mathbf{u})$, the system (1.1) implies an additional conservation law for the total energy density defined as

$$E(r, u) := \tfrac{1}{2} r|u|^2 + U(r) \quad \text{for all } (r, u) \in \mathbb{U}.$$

In fact, we have that

$$\partial_t \Big(\tfrac{1}{2} \varrho |\mathbf{u}|^2 + U(\varrho) \Big) + \nabla \cdot \bigg(\Big(\tfrac{1}{2} \varrho |\mathbf{u}|^2 + Q(\varrho) \Big) \mathbf{u} \bigg) = 0 \quad \text{in } [0, \infty) \times \mathbf{R}^d, \tag{1.4}$$

where the function Q can be expressed in terms of the internal energy as

$$Q(r) = U'(r) r \quad \text{for all } r \geqslant 0.$$

Equation (1.4) implies that for smooth solutions $(\varrho, \mathbf{u})$ the total energy is conserved. A typical phenomenon in compressible fluid dynamics, however, is the occurrence of jump discontinuities (shocks) in the conserved quantities, so (1.4) cannot hold globally. Since (1.1) describes a closed system, in which no energy is injected from outside, a reasonable relaxation of (1.4) is to require an inequality

$$\partial_t \Big(\tfrac{1}{2} \varrho |\mathbf{u}|^2 + U(\varrho) \Big) + \nabla \cdot \bigg(\Big(\tfrac{1}{2} \varrho |\mathbf{u}|^2 + Q(\varrho) \Big) \mathbf{u} \bigg) \leqslant 0 \tag{1.5}$$

in the sense of distributions. This inequality implies that energy can only decrease (be transformed into different forms of energy not represented in the model, such as heat), but no energy can be created. The total energy E is convex in the conserved quantities $(\varrho, \varrho\mathbf{u})$. As a consequence of (1.1) and (1.5), one obtains that

(1) **The Total Mass is Conserved.**
For all $t \in [0, \infty)$ we have

$$\int_{\mathbf{R}^d} \varrho(t, x)\, dx = \int_{\mathbf{R}^d} \bar{\varrho}(x)\, dx. \tag{1.6}$$

(2) **The Total Energy is Nonincreasing in Time.**
For a.e. $0 \leqslant t_1 \leqslant t_2 < \infty$ we have

$$\int_{\mathbf{R}^d} \Big(\tfrac{1}{2} \varrho |\mathbf{u}|^2 + U(\varrho) \Big)(t_2, x)\, dx \leqslant \int_{\mathbf{R}^d} \Big(\tfrac{1}{2} \varrho |\mathbf{u}|^2 + U(\varrho) \Big)(t_1, x)\, dx. \tag{1.7}$$

We assume that the initial data $(\bar{\varrho}, \bar{\mathbf{u}})$ has finite total mass and total energy. Then (1.6) and (1.7) imply that the mass and total energy stay bounded for all time.

Pairs of functions $(\eta, q) \colon \mathbb{U} \longrightarrow \mathbf{R} \times \mathbf{R}^d$ with the property that

$$\partial_t \eta(\varrho, \mathbf{u}) + \nabla \cdot q(\varrho, \mathbf{u}) = 0 \tag{1.8}$$

in all open sets in $[0, \infty) \times \mathbf{R}^d$ where the solution $(\varrho, \mathbf{u})$ of (1.1) is smooth, are called entropy/entropy-flux pairs. Due to the occurence of shocks, equation (1.8) cannot hold globally, so one typically requires a differential inequality in the sense of distributions, for all entropy/entropy-flux (η, q) such that η is convex in the conserved quantities $(\varrho, \varrho\mathbf{u})$. This assumption is called an entropy condition.

In one space dimension, there exists a large class of convex entropy/entropy-flux pairs. They provide a priori estimates that are crucial for all global existence results that are available for this case; see [16] and the references therein. In the multidimensional case, the only entropy seems to be the total energy. Dafermos [9] suggested to strengthen the entropy condition for conservation laws by looking not for any weak solution that satisfies the entropy inequalities (like the one in (1.5) for solutions of the isentropic Euler equations (1.1)), but instead to look for that particular solution for which the total entropy (total energy in the case of (1.1)) decreases as fast as possible. Such a steepest descent approach has already been very successful in the context of certain degenerate parabolic equations. It was first developed by Otto [22] for the porous medium equation and relies on the theory of optimal transport and the calculus of variations.

To explain the approach and to introduce some notation, let us explain Otto's result in some detail. Let P and U be the pressure and internal energy introduced above, satisfying the relation (1.2). Then the porous medium equation

$$\partial_t \varrho - \Delta P(\varrho) = 0 \quad \text{in } [0,\infty) \times \mathbf{R}^d \tag{1.9}$$

can be considered as a gradient flow on a suitable manifold, in the sense that the solutions of (1.9) are precisely those curves along which the internal energy

$$\mathcal{U}(\varrho) := \int_{\mathbf{R}^d} U(\varrho)\, dx \tag{1.10}$$

decreases with maximal rate. Recall that for any gradient flow $\dot{X} = -\nabla E(X)$, the following mathematical structure must be given: we need

(1) a manifold $\mathcal{M}$ with tangent space $\mathbb{T}_X\mathcal{M}$ for $X \in \mathcal{M}$,
(2) a smooth function $E\colon \mathcal{M} \longrightarrow \mathbf{R}$ with differential $dE(X) \in \mathbb{T}_X\mathcal{M}^*$, and
(3) a metric tensor g, which allows us to define $\nabla E(X) \in \mathbb{T}_X\mathcal{M}$ by

$$g_X\big(\nabla E(X), V\big) := dE(X)V \quad \text{for all } V \in \mathbb{T}_X\mathcal{M}.$$

For the porous medium equation these objects are given as follows:

(1) The manifold is the space $\mathscr{P}(\mathbf{R}^d)$ of probability measures with finite second moments, that are absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^d$. This space is equipped with the Wasserstein distance, which is defined by

$$\mathbf{W}(\varrho_1, \varrho_2)^2 := \inf\left\{ \iint_{\mathbf{R}^d\times\mathbf{R}^d} |x_2 - x_1|^2\, \gamma(dx_1, dx_2)\colon \pi^i\#\gamma = \varrho_i\mathcal{L}^d \right\} \tag{1.11}$$

for all $\varrho_1, \varrho_2 \in \mathscr{P}(\mathbf{R}^d)$. The number in (1.11) is the minimal quadratic cost required to transport the measure $\varrho_1\mathcal{L}^d$ to the measure $\varrho_2\mathcal{L}^d$. The map $\pi^i\colon \mathbf{R}^d\times\mathbf{R}^d \longrightarrow \mathbf{R}^d$ is the projection onto the ith component, and $\#$ denotes the pushforward. The probability measure γ on $\mathbf{R}^d\times\mathbf{R}^d$ is called a transport plan, and one can show that the inf in (1.11) is attained. In fact, the optimal transport plan is induced by a map: there exists a convex function $\phi\colon \mathbf{R}^d \longrightarrow \mathbf{R}$ such that $\gamma = (\mathrm{id}, \nabla\phi)_{\#}\varrho_1$.

(2) For any $\varrho \in \mathscr{P}(\mathbf{R}^d)$, we now define the tangent space $\mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$ as

$$\text{the closure of } \Big\{\nabla\phi\colon \phi \in \mathscr{D}(\mathbf{R}^d)\Big\} \text{ in the } \mathscr{L}^2(\mathbf{R}^d, \varrho)\text{-norm.}$$

Then the $\mathscr{L}^2(\mathbf{R}^d, \varrho)$-inner product induces a metric on $\mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$. This definition is motivated by the fact that for any sufficiently smooth curve $t \mapsto \varrho(t) \in \mathscr{P}(\mathbf{R}^d)$ with $\varrho(0) = \varrho$, there exists a unique $\mathbf{u} \in \mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$ with the property that

$$\partial_t\varrho + \nabla\cdot(\varrho\mathbf{u}) = 0 \quad \text{in } \mathscr{D}'(\mathbf{R}^d). \tag{1.12}$$

There is in fact an orthogonal decomposition (see [5])

$$\mathscr{L}^2(\mathbf{R}^d, \varrho) = \mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d) \oplus \mathbb{N}_\varrho\mathscr{P}(\mathbf{R}^d), \tag{1.13}$$

where the normal space is the space of divergence-free vector fields:

$$\mathbb{N}_\varrho \mathscr{P}(\mathbf{R}^d) := \Big\{ \mathbf{w} \in \mathscr{L}^2(\mathbf{R}^d, \varrho) \colon \nabla \cdot (\varrho \mathbf{w}) = 0 \text{ in } \mathscr{D}'(\mathbf{R}^d) \Big\}. \tag{1.14}$$

(3) The internal energy (1.10) is displacement convex (that is, convex along the geodesics of the metric space $(\mathscr{P}(\mathbf{R}^d), \mathbf{W})$). Its subdifferential is defined as

$$\partial \mathcal{U}(\varrho) := \Big\{ \mathbf{v} \in \mathbb{T}_\varrho \mathscr{P}(\mathbf{R}^d) \colon \\ \mathcal{U}(\varrho^*) \geqslant \mathcal{U}(\varrho) + \int_{\mathbf{R}^d} \mathbf{v} \cdot (\mathbf{r} - \mathrm{id})\, \varrho \text{ for all } \mathbf{r} \in \mathrm{Opt}(\varrho, \varrho^*) \Big\},$$

and the unique element in $\partial \mathcal{U}(\varrho)$ with minimal $\mathscr{L}^2(\mathbf{R}^d, \varrho)$-norm is given by

$$\nabla U'(\varrho) = \varrho^{-1} \nabla P(\varrho).$$

If we now use $\mathbf{u} := -\varrho^{-1} \nabla P(\varrho)$ in the continuity equation (1.12), then we obtain the porous medium equation (1.9) at one instance in time.

This result has been generalized considerably, and we refer the reader to the recent monographs [5,26] for more details. The interpretation of dissipative evolution equations as abstract gradient flows on the Wasserstein space $\mathscr{P}(\mathbf{R}^d)$ suggests a natural time discretization: Given a timestep $\tau > 0$ and the approximate density $\varrho^n \in \mathscr{P}(\mathbf{R}^d)$ at time $t^n := n\tau$, the value at time t^{n+1} is defined as

$$\varrho^{n+1} \in \operatorname{argmin} \left\{ \frac{\mathbf{W}(\varrho^n, \varrho)^2}{2\tau} + \mathcal{U}(\varrho) \colon \varrho \in \mathscr{P}(\mathbf{R}^d) \right\}; \tag{1.15}$$

see [14]. One can show that this approximation converges to a solution of (1.9) as $\tau \to 0$. Since the multidimensional isentropic Euler equations are not a gradient flow, the above framework cannot be used. There is considerable interest recently to develop an analogous theory for Hamiltonian systems; see e.g. [3].

In [12] we proposed a variational time discretization similar to (1.15) for the isentropic Euler equations (1.1). For initial data $(\bar{\varrho}, \bar{\mathbf{u}}) =: (\varrho^0, \mathbf{u}^0)$ with finite energy and given timestep $\tau > 0$, we construct approximate solutions $(\varrho^n, \mathbf{u}^n)$ of (1.1) at times $t^n := n\tau$ with $n \in \mathbf{N}$ by iterating the three steps of the following scheme:

(1) **Velocity Projection.**
(2) **Free Transport.**
(3) **Energy Minimization.**

In the velocity projection step, we substitute for $\mathbf{u}^n$ a velocity $\hat{\mathbf{u}}^n \in \mathscr{L}^2(\mathbf{R}^d, \varrho^n)$ that has the property that the vector field $\hat{\mathbf{r}}^n := \mathrm{id} + \tau \hat{\mathbf{u}}^n$ is an optimal tranport map (a gradient of a convex function). In [12] we used Brenier's polar factorization [8] to construct the new velocity $\hat{\mathbf{u}}^n$ from the old one $\mathbf{u}^n$. In this paper, we propose to compute the metric projection onto the cone of optimal transport maps instead. We refer the reader to Section 2, where we will explain this step in detail.

In the free transport step, we push $\varrho^n \mathcal{L}^d$ forward along the optimal transport map $\hat{\mathbf{r}}^n = \mathrm{id} + \tau \hat{\mathbf{u}}^n$, which defines a new measure $\hat{\varrho}^n := \hat{\mathbf{r}}^n_\# (\varrho^n \mathcal{L}^d)$.

In the energy minimization step, we then compute

$$\varrho^{n+1} \in \operatorname{argmin} \left\{ \frac{3}{4\tau^2} \mathbf{W}(\hat{\varrho}^n, \varrho)^2 + \mathcal{U}(\varrho) \colon \varrho \in \mathscr{P}(\mathbf{R}^d) \right\}.$$

This minimization is similar to the one in (1.15) for the porous medium equation. There exists a unique map $\mathbf{r}^{n+1} \in \operatorname{Opt}(\varrho^{n+1}, \hat{\varrho}^n)$, which is given by

$$\mathbf{r}^{n+1} := \mathrm{id} + \frac{2\tau^2}{3} \nabla U'(\varrho^{n+1}); \tag{1.16}$$

see [5]. The new density/velocity $(\varrho^{n+1}, \mathbf{u}^{n+1})$ then satisfy the identities

$$\varrho^{n+1}\mathcal{L}^d = \left(\Big(\mathrm{id} + \tau\hat{\mathbf{u}}^n\Big)^{-1} \circ \Big(\mathrm{id} + \frac{2\tau^2}{3}\nabla U'(\varrho^{n+1})\Big) \right)^{-1}_{\#} (\varrho^n \mathcal{L}^d), \tag{1.17}$$

$$\mathbf{u}^{n+1} = \hat{\mathbf{u}}^n \circ \Big(\mathrm{id} + \tau\hat{\mathbf{u}}^n\Big)^{-1} \circ \Big(\mathrm{id} + \frac{2\tau^2}{3}\nabla U'(\varrho^{n+1})\Big) - \tau \nabla U'(\varrho^{n+1}). \tag{1.18}$$

They can be justified as the minimizers of an optimization problem for the density/velocity similar to (1.15), where the Wasserstein distance is replaced by a new cost functional called the Minimal Acceleration Cost, which measures the deviation of particles from their free transport paths. The intuition here is that fluid particles prefer to travel on straight lines in the direction of their velocity, but may deviate if by doing so, the internal energy can be decreased; see [12].

Since (1.16) is a second order in τ perturbation of the identity, equation (1.17) indicates that up to second order, the new density ϱ^{n+1} is obtained by transporting ϱ^n in the direction of the optimal velocity field $\hat{\mathbf{u}}^n$. This is the discrete analogue of the continuity equation (1.12). Similarly, the new velocity $\mathbf{u}^{n+1}$ is up to second order equal to the old velocity $\hat{\mathbf{u}}^n$ transported in the direction of $\hat{\mathbf{u}}^n$, minus the pressure term $\nabla U'(\varrho^{n+1})$. This corresponds to the velocity equation

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla U'(\varrho) = 0, \tag{1.19}$$

which formally follows from the continuity and the momentum equation in (1.1).

The variational time discretization is well-defined (see [12] and Section 2 below). Moreover, in each step the total energy is decreased: If we define

$$\mathcal{E}(\varrho, \mathbf{u}) := \int_{\mathbf{R}^d} \Big(\tfrac{1}{2}\varrho|\mathbf{u}|^2 + U(\varrho) \Big)(x)\, dx$$

for all densities/velocities $(\varrho, \mathbf{u})$, then we have the energy dissipation estimate

$$\mathcal{E}(\varrho^{n+1}, \mathbf{u}^{n+1}) + \frac{\tau^2}{6} \int_{\mathbf{R}^d} |\nabla U'(\varrho^{n+1})|^2 \varrho^{n+1}\, dx \leqslant \mathcal{E}(\varrho^n, \mathbf{u}^n). \tag{1.20}$$

The second integral on the left-hand side of (1.20) allows us to control the second order terms that appear in equations (1.17) and (1.18). We proved in [12] that a sequence of approximate solutions generated by our variational time discretization converges weakly to a density/velocity $(\varrho, \mathbf{u})$, which satisfy the continuity equation (1.12). We will consider the momentum equation in Section 2.

We explained above that the porous medium equation can be considered as a gradient flow on the space of probability measures $\mathscr{P}(\mathbf{R}^d)$, and thus as a first-order ordinary differential equation. Here the velocity is given by an "algebraic" relation for the velocity $\mathbf{u}$, which is Darcy's law $\mathbf{u} = -\nabla U'(\varrho)$. It only involves the value of the density at a given time. In contrast, for the isentropic Euler equations the velocity or momentum equations involve the time derivative of $\mathbf{u}$, as in (1.19). If we think of $\mathbf{u}$ as a tangent vector on the manifold $\mathscr{P}(\mathbf{R}^d)$, then (1.19) must have an interpretation in terms of *covariant* derivatives and we need a second-order calculus for $\mathscr{P}(\mathbf{R}^d)$, which has attracted a lot of interest recently. In the following, we will review the basics of this theory and show how the isentropic Euler equations fit into this framework as a second order differential equations.

As is done in classical differential geometry, let us consider a curve in our manifold, which is a one-parameter family of probability measures $t \mapsto \varrho_t \in \mathscr{P}(\mathbf{R}^d)$. We assume that a tangent vector field $t \mapsto \mathbf{v}_t \in \mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d)$ is defined along the curve. In order to differentiate $\mathbf{v}_t$ along the curve, we first need a way to transport vectors from one tangent space at ϱ_s to another tangent space at ϱ_t. Notice that since $\mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d) \subset \mathscr{L}^2(\mathbf{R}^d, \varrho_t)$, these spaces may be very different for $s \neq t$. For simplicity of presentation, let us assume that the curve $t \mapsto \varrho_t$ admits a velocity vector $\mathbf{u}_t \in \mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d)$ and that the vector field $\mathbf{u}_t$ is so smooth that the flow

$$\dot{X}_t(s) = \mathbf{u}_s \circ X_t(s), \quad X_t(t) = \mathrm{id} \quad \text{for all } s, t$$

is a diffeomorphism. Then the pull-back under the flow, given by

$$X_t(s)(\mathbf{v}_s) := \mathbf{v}_s \circ X_t(s) \quad \text{for all } s, t,$$

is an isometry from $\mathscr{L}^2(\mathbf{R}^d, \varrho_s)$ to $\mathscr{L}^2(\mathbf{R}^d, \varrho_t)$. We now define the total derivative

$$\frac{D}{dt}\mathbf{v}_t := \lim_{s \to t} \frac{X_t(s)(\mathbf{v}_s) - \mathbf{v}_t}{s - t} \quad \text{in } \mathscr{L}^2(\mathbf{R}^d, \varrho_t). \tag{1.21}$$

Even if the vector field is tangent along the curve: that is, if $\mathbf{v}_t \in \mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d)$ for all t, the total derivative $\frac{D}{dt}\mathbf{v}_t$ in general is not a tangent vector. In order to restore the tangency, we combine the total derivative with a projection: Recall that $\mathscr{L}^2(\mathbf{R}^d, \varrho)$ admits an orthogonal decomposition into the space of tangent and normal vector fields; see (1.13). We define the orthogonal projection

$$P_{\varrho_t} \colon \mathscr{L}^2(\mathbf{R}^d, \varrho_t) \longrightarrow \mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d).$$

Then the covariant derivative of $\mathbf{v}_t$ in the direction of $\mathbf{u}_t$ is given by

$$\nabla_{\mathbf{u}_t}\mathbf{v}_t := P_{\varrho_t}\left(\frac{D}{dt}\mathbf{v}_t\right) = P_{\varrho_t}\Big(\partial_t \mathbf{v}_t + \mathbf{u}_t \cdot \nabla \mathbf{v}_t\Big). \tag{1.22}$$

The last identity follows from the fact that the computation of the total derivative in (1.21) requires both a partial derivative with respect to the parameter t as well as a convective derivative, which takes into account the action of the flow. One can show that definition (1.22) is natural in the sense that it is the analogue of the Levi-Civita connection and therefore compatible with the inner product on $\mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d)$. We refer the reader to [4, 13, 18] for further information.

We now compute the covariant derivative of the velocity $t \mapsto \mathbf{u}_t \in \mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d)$ in the direction of the flow (in the direction of $\mathbf{u}_t$). Recall that tangent vectors are (limits of) gradient vector fields. Assuming that formally we can write $\mathbf{u}_t = \nabla\phi_t$ for a suitable potential field ϕ_t, we then obtain the following identity

$$\nabla_{\mathbf{u}_t}\mathbf{u}_t = P_{\varrho_t}\Big(\partial_t \nabla\phi_t + D^2\phi_t \cdot \nabla\phi_t\Big) = \nabla\Big(\partial_t\phi_t + \frac{1}{2}|\nabla\phi_t|^2\Big). \tag{1.23}$$

The left-hand side represents the second derivative of the curve $t \mapsto \varrho_t$.

We now invoke Newton's law, which states that an acceleration must be compensated by a force. For the isentropic Euler equations the force comes from the pressure due to the presence of nearby fluid elements. We thus consider

$$\nabla_{\mathbf{u}_t}\mathbf{u}_t = -\nabla U'(\varrho_t). \tag{1.24}$$

Since $\nabla U'(\varrho_t) \in \mathbb{T}_{\varrho_t}\mathscr{P}(\mathbf{R}^d)$ the projection P_{ϱ_t} leaves the pressure term unchanged. We combine (1.23) with (1.24) to obtain the modified velocity equation

$$P_{\varrho_t}\bigg(\partial_t \mathbf{u}_t + \mathbf{u}_t \cdot \nabla \mathbf{u}_t + \nabla U'(\varrho_t)\bigg) = 0. \tag{1.25}$$

If we again write $\mathbf{u}_t = \nabla\phi_t$ for a suitable potential ψ_t, then

$$\nabla\bigg(\partial_t\phi_t + \frac{1}{2}|\nabla\phi_t|^2 + U'(\varrho_t)\bigg) = 0 \quad \text{in spt}\,\varrho_t,$$

which is Bernoulli's law. The second-order ordinary differential equation (1.24) on the Wasserstein space $\mathscr{P}(\mathbf{R}^d)$ therefore represents the isentropic Euler equations in the regime of potential flows, for which the velocity is a gradient vector field. The continuity equation (1.12) plays the role of a compatibility condition for the density and the velocity. Among all velocity fields in $\mathscr{L}^2(\mathbf{R}^d, \varrho_t)$ that satisfy (1.12), the gradient vector fields are those with minimal kinetic energy.

In this paper, we will make some of the formal arguments above more rigorous. As usual in the theory of conservation laws, it will be important that the equations we consider are in conservative form. In order to understand what equation (1.25) implies, note first that in our discussion above we assumed that

$$R_t := \partial_t \mathbf{u}_t + \mathbf{u}_t \cdot \nabla \mathbf{u}_t + \nabla U'(\varrho_t) \in \mathscr{L}^2(\mathbf{R}^d, \varrho_t).$$

We will show below that such a statement is true on the level of the discretization. By (1.25), the tangent component of R_t vanishes and so the vector field R_t must be normal to the manifold $\mathscr{P}(\mathbf{R}^d)$; see the orthogonal decomposition (1.13). According to the characterization (1.14), this entails that R_t is divergence-free:

$$\nabla\cdot\bigg(\varrho_t\Big(\partial_t \mathbf{u}_t + \mathbf{u}_t \cdot \nabla \mathbf{u}_t + \nabla U'(\varrho_t)\Big)\bigg) = 0.$$

Using the continuity equation (1.12), we obtain that $(\varrho_t, \mathbf{u}_t)$ satisfy

$$\nabla\cdot\bigg(\partial_t(\varrho_t\mathbf{u}_t) + \nabla\cdot(\varrho_t\mathbf{u}_t \otimes \mathbf{u}_t) + \nabla P(\varrho_t)\bigg) = 0$$

in distributional sense. We therefore only test against gradient vector fields instead of against all $\mathbf{R}^d$-valued test functions $\zeta \in \mathscr{D}(\mathbf{R}^d)$, and so the interpretation of the system of isentropic Euler equations as second order ordinary differential equation on $\mathscr{P}(\mathbf{R}^d)$ only implies a modified momentum equation. The full information from (1.1) is restored when we take into account the fact that velocity fields are supposed to be tangent and therefore gradient vector fields.

Here is an outline of our paper: In Section 2, we will discuss the metric projection of vector fields in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$ onto the closed and convex cone of optimal transport maps, which are gradients of convex functions. In Section 3, we will show that the difference between a given vector field and its metric projection can be expressed in terms of a matrix-valued measure, which we call a stress tensor field. We will derive bounds on the stress tensor. In Section 5, we will consider a sequence of approximate solutions to (1.1) that are obtained from our variational time discretization in the limit $\tau \to 0$. We will show that a suitable subsequence converges weakly, and in fact generates a Young measure that captures possible oscillations. In Section 6, we will show that this Young measure is a measure-valued solution to (1.1).

2. Velocity Projection

In the following, we will denote by $\mathscr{P}(\mathbf{R}^d)$ the space of probability measures on $\mathbf{R}^d$ with finite second moments. We will use the short-hand notation $\varrho \ll \mathcal{L}^d$ to indicate when a measure ϱ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^d$. To simplify the notation, we will not distinguish between a density ϱ and its induced measure $\varrho\mathcal{L}^d$ whenever the meaning is clear from the context.

For any $\varrho \in \mathscr{P}(\mathbf{R}^d)$ we define the tangent bundle

$$\mathbb{T}\mathscr{P}(\mathbf{R}^d) := \Big\{(\varrho, \mathbf{u})\colon \varrho \in \mathscr{P}(\mathbf{R}^d), \mathbf{u} \in \mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)\Big\},$$

with tangent spaces $\mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$ given as

$$\text{the closure of } \Big\{\nabla\varphi\colon \varphi \in \mathscr{D}(\mathbf{R}^d)\Big\} \text{ in the } \mathscr{L}^2(\mathbf{R}^d, \varrho)\text{-norm.}$$

We will occasionally write $\langle a, b\rangle := a \cdot b$ for all $a, b \in \mathbf{R}^d$ for the Euclidean inner product, and we will denote by $\pi^k\colon \mathbf{R}^{Nd} \longrightarrow \mathbf{R}^d$ the projection of $\mathbf{R}^{Nd} = (\mathbf{R}^d)^N$ onto its $(k+1)$st factor, where $N \in \mathbf{N}$ and $k = 0 \dots N-1$.

Definition 2.1 (Distance)**.** Let $\varrho \in \mathscr{P}(\mathbf{R}^d)$ be given. Then we define

$$\mathscr{P}_\varrho(\mathbf{R}^{2d}) := \Big\{\boldsymbol{\gamma} \in \mathscr{P}(\mathbf{R}^{2d})\colon \pi^0_\#\boldsymbol{\gamma} = \varrho\Big\}.$$

We introduce a distance as follows: For any $\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2 \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ we let

$$\mathbf{W}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)^2 := \int_{\mathbf{R}^d} \mathbf{W}(\gamma^1_x, \gamma^2_x)^2\, \varrho(dx),$$

where $\boldsymbol{\gamma}^k(dx, dy) =: \gamma^k_x(dy)\, \varrho(dx)$ with $k = 1 \dots 2$ denotes the disintegration of $\boldsymbol{\gamma}^k$ with respect to ϱ. Here $\mathbf{W}$ is the Wasserstein distance defined as

$$\mathbf{W}(\varrho_1, \varrho_2)^2 := \inf\left\{\iint_{\mathbf{R}^d\times\mathbf{R}^d} |x_2 - x_1|^2\, \gamma(dx_1, dx_2)\colon \pi^i\#\gamma = \varrho_i\right\}$$

for all $\varrho_1, \varrho_2 \in \mathscr{P}(\mathbf{R}^d)$. The inf is attained for an optimal transport plan γ, which is supported in the subdifferential of a lower semicontinuous, convex function. We denote by $\mathrm{Opt}(\varrho_1, \varrho_2)$ the set of optimal transport plans from ϱ_1 to ϱ_2.

Definition 2.2 (Transport Plans)**.** Let $\varrho \in \mathscr{P}(\mathbf{R}^d)$ be given.

(i.) *Admissible Plans.* For any $\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2 \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ we define

$$\mathrm{Adm}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2) := \Big\{\boldsymbol{\alpha} \in \mathscr{P}(\mathbf{R}^{3d})\colon (\pi^0, \pi^1)_\#\boldsymbol{\alpha} = \boldsymbol{\gamma}^1, (\pi^0, \pi^2)_\#\boldsymbol{\alpha} = \boldsymbol{\gamma}^2\Big\}.$$

(ii.) *Optimal Plans.* For any $\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2 \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ we define

$$\mathrm{Opt}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2) := \Big\{\boldsymbol{\alpha} \in \mathrm{Adm}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)\colon \mathbf{W}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)^2 = \int_{\mathbf{R}^{3d}} |y^1 - y^2|^2\, \boldsymbol{\alpha}(dx, dy^1, dy^2)\Big\}.$$

Theorem 2.3. *Let $\varrho \in \mathscr{P}(\mathbf{R}^d)$ be given.*

(i.) *The function $\mathbf{W}_\varrho$ is a distance on $\mathscr{P}_\varrho(\mathbf{R}^{2d})$ and lower semicontinuous with respect to the weak convergence in $\mathscr{P}(\mathbf{R}^{2d})$. We have*

$$\mathbf{W}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)^2 = \min_{\boldsymbol{\alpha}\in\mathrm{Adm}_\varrho(\boldsymbol{\gamma}^1,\boldsymbol{\gamma}^2)} \int_{\mathbf{R}^{3d}} |y^1 - y^2|^2\, \boldsymbol{\alpha}(dx, dy^1, dy^2)$$

for all $\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2 \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$. In particular, the set $\mathrm{Opt}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)$ is nonempty.

(ii.) *The set* $(\mathscr{P}_\varrho(\mathbf{R}^{2d}), \mathbf{W}_\varrho)$ *is a complete metric space.*

Proof. We refer the reader to Section 4.1 in [13]. ☐

Remark 2.4. An important subset of $\mathscr{P}_\varrho(\mathbf{R}^{2d})$ consists of those measures $\boldsymbol{\gamma}$ that are induced by maps: there exists a function $\mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ such that $\boldsymbol{\gamma} = (\mathrm{id}, \mathbf{r})_\# \varrho$. Note that if $\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2 \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ and $\boldsymbol{\gamma}^1$ is induced by a map $\mathbf{r}^1$, then

$$\mathbf{W}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)^2 = \int_{\mathbf{R}^d} |\mathbf{r}^1(x) - y|^2 \, \boldsymbol{\gamma}^2(dx, dy),$$

as follows from (5.2.12) in [5]. If in addition $\boldsymbol{\gamma}^2$ is induced by a map $\mathbf{r}^2$, then

$$\mathbf{W}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)^2 = \int_{\mathbf{R}^d} |\mathbf{r}^1(x) - \mathbf{r}^2(x)|^2 \, \varrho(dx),$$

which is the usual $\mathscr{L}^2(\mathbf{R}^d, \varrho)$-distance between $\mathbf{r}^1$ and $\mathbf{r}^2$. We thereby identify the Hilbert space $\mathscr{L}^2(\mathbf{R}^d, \varrho)$ with a closed subset of $(\mathscr{P}_\varrho(\mathbf{R}^{2d}), \mathbf{W}_\varrho)$.

Consider now the set of optimal transport maps

$$\mathscr{C}_\varrho := \Big\{ \mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d, \varrho) \colon (\mathrm{id}, \mathbf{r})_\# \varrho \in \mathrm{Opt}(\varrho, \mathbf{r}_\# \varrho) \Big\} \tag{2.1}$$

and the set of induced optimal transport plans $C_\varrho := \{(\mathrm{id}, \mathbf{r})_\# \varrho \colon \mathbf{r} \in \mathscr{C}_\varrho\}$. Let us also introduce the isomorphism $\iota : \mathscr{C}_\varrho \longrightarrow C_\varrho$ defined by

$$\iota(\mathbf{r}) := (\mathrm{id}, \mathbf{r})_\# \varrho \quad \text{for all } \mathbf{r} \in \mathscr{C}_\varrho. \tag{2.2}$$

Using Remark 2.4, we find that for all $\mathbf{r}^1, \mathbf{r}^2 \in \mathscr{C}_\varrho$ we have

$$\mathbf{W}_\varrho\Big(\iota(\mathbf{r}^1), \iota(\mathbf{r}^2)\Big) = \|\mathbf{r}^1 - \mathbf{r}^2\|_{\mathscr{L}^2(\mathbf{R}^d, \varrho)},$$

so the isomorphism (2.2) is an isometry if $\mathscr{C}_\varrho$ is equipped with the usual $\mathscr{L}^2(\mathbf{R}^d, \varrho)$-distance, and C_ϱ is equipped with the distance $\mathbf{W}_\varrho$ of Definition 2.1.

Remark 2.5. We emphasize the fact that we only consider optimal transport *maps*, not plans. This is no restriction if the measure ϱ is absolutely continuous with respect to the Lebesgue measure. Our choice is motivated by the connection between the metric projection onto $\mathscr{C}_\varrho$ and the sticky particle dynamics for pressureless gas dynamics. We refer the reader to Section 5 for further information.

Proposition 2.6. *Let* $\varrho \in \mathscr{P}(\mathbf{R}^d)$ *be given and let* $\mathscr{C}_\varrho$ *be the set of optimal transport maps defined by* (2.1). *Then* $\mathscr{C}_\varrho$ *is a closed, convex cone in* $\mathscr{L}^2(\mathbf{R}^d, \varrho)$.

Proof. Consider first a sequence of optimal transport maps $\mathbf{r}^n \in \mathscr{C}_\varrho$ with $\mathbf{r}^n \longrightarrow \mathbf{r}$ for some $\mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$. Then the induced optimal transport plans $\boldsymbol{\gamma}^n := (\mathrm{id}, \mathbf{r}^n)_\# \varrho$ converge narrowly to the plan $\boldsymbol{\gamma} := (\mathrm{id}, \mathbf{r})_\# \varrho \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$; see Lemma 5.4.1 in [5]. But the narrow limit of a sequence of optimal transport plans is again an optimal transport plan (see Proposition 7.1.3 in [5]), and therefore $\mathbf{r} \in \mathscr{C}_\varrho$.

Recall now that a transport plan $\boldsymbol{\gamma} \in \mathscr{P}(\mathbf{R}^{2d})$ is optimal if and only if its support is cyclically monotone; see Section 6.2.3 in [5]. That is, if we have

$$\sum_{i=1}^N \langle x^i - x^{i-1}, y^i \rangle \geqslant 0 \quad \begin{array}{l} \text{for any } N \in \mathbf{N} \text{ and any choice of } (x^i, y^i) \in \mathrm{spt}\, \boldsymbol{\gamma} \\ \text{with } i = 1 \ldots N, \text{ where } x^0 := x^N. \end{array} \tag{2.3}$$

Notice that the plan $\boldsymbol{\gamma}_b := (\mathrm{id}, b)_\# \varrho$, which pushes ϱ forward to the Dirac measure located at position $b \in \mathbf{R}^d$, trivially satisfies (2.3). Hence $\mathscr{C}_\varrho$ contains all constant functions. Consider now the optimal transport plan $\boldsymbol{\gamma} = (\mathrm{id}, \mathbf{r})_\# \varrho$ for some $\mathbf{r} \in \mathscr{C}_\varrho$.

Define $\boldsymbol{\gamma}^s := (\mathrm{id}, s\mathbf{r})_\# \varrho$ with $s > 0$. Then $(x,y) \in \operatorname{spt} \boldsymbol{\gamma}^s$ if and only if $y = s\mathbf{r}(x)$, which in turn is equivalent to $(x, y/s) \in \operatorname{spt} \boldsymbol{\gamma}$. Since

$$\sum_{i=1}^N \langle x^i - x^{i-1}, y^i \rangle = s \sum_{i=1}^N \langle x^i - x^{i-1}, y^i/s \rangle$$

for any $(x^i, y^i) \in \operatorname{spt} \boldsymbol{\gamma}^s$, and since $\operatorname{spt} \boldsymbol{\gamma}$ is cyclically monotone, we obtain (2.3).

Similarly, if $\boldsymbol{\gamma}^k = (\mathrm{id}, \mathbf{r}^k)_\# \varrho$ for $\mathbf{r}^k \in \mathscr{C}_\varrho$ and $k = 1 \ldots 2$, then we define

$$\boldsymbol{\gamma}^{[s]} := \Big(\mathrm{id}, (1-s)\mathbf{r}^1 + s\mathbf{r}^2\Big)_\# \varrho \quad \text{for } s \in [0,1].$$

Then $(x,y) \in \operatorname{spt} \boldsymbol{\gamma}^{[s]}$ if and only if $y = (1-s)\mathbf{r}^1(x) + s\mathbf{r}^2(x)$. On the other hand, defining $y^k := \mathbf{r}^k(x)$ we have that $(x, y^k) \in \operatorname{spt} \boldsymbol{\gamma}^k$, for $k = 1 \ldots 2$. Since

$$\sum_{i=1}^N \langle x^i - x^{i-1}, y^i \rangle = (1-s) \sum_{i=1}^N \langle x^i - x^{i-1}, y^{i,1} \rangle + s \sum_{i=1}^N \langle x^i - x^{i-1}, y^{i,2} \rangle$$

for any $(x^i, y^i) \in \operatorname{spt} \boldsymbol{\gamma}^{[s]}$ and $y^{i,k} := \mathbf{r}^k(x^i)$, and since $\operatorname{spt} \boldsymbol{\gamma}^k$ is cyclically monotone, we obtain (2.3). This proves that $\mathscr{C}_\varrho$ is a closed, convex cone in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$. □

For any given transport plan $\boldsymbol{\gamma} \in \mathscr{P}_\varrho(\mathbf{R}^d)$ we want to find an element in C_ϱ that is closest to $\boldsymbol{\gamma}$ with respect to the distance $\mathbf{W}_\varrho$. Since any plan $\boldsymbol{\eta} \in C_\varrho$ is induced by an optimal transport map $\mathbf{r} \in \mathscr{C}_\varrho$, Remark 2.4 implies that

$$\mathbf{W}_\varrho(\boldsymbol{\gamma}, \boldsymbol{\eta})^2 = \int_{\mathbf{R}^{2d}} |y - \mathbf{r}(x)|^2 \, \boldsymbol{\gamma}(dx, dy).$$

Theorem 2.7. *Let $\varrho \in \mathscr{P}(\mathbf{R}^d)$ be given and let $\mathscr{C}_\varrho$ be the set of optimal transport maps defined by (2.1). For all $\boldsymbol{\gamma} \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ there is a unique $\mathbf{r}_{\boldsymbol{\gamma}} \in \mathscr{C}_\varrho$ with*

$$\int_{\mathbf{R}^{2d}} |y - \mathbf{r}_{\boldsymbol{\gamma}}(x)|^2 \, \boldsymbol{\gamma}(dx, dy) \leqslant \int_{\mathbf{R}^{2d}} |y - \mathbf{r}(x)|^2 \, \boldsymbol{\gamma}(dx, dy) \quad \textit{for all } \mathbf{r} \in \mathscr{C}_\varrho. \tag{2.4}$$

The minimizer $\mathbf{r}_{\boldsymbol{\gamma}}$ has the following properties:

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbf{r}_{\boldsymbol{\gamma}}(x), \mathbf{r}_{\boldsymbol{\gamma}}(x) \rangle \, \boldsymbol{\gamma}(dx, dy) = 0, \tag{2.5}$$

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbf{r}_{\boldsymbol{\gamma}}(x), \mathbf{r}_{\boldsymbol{\gamma}}(x) - x \rangle \, \boldsymbol{\gamma}(dx, dy) \geqslant 0, \tag{2.6}$$

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbf{r}_{\boldsymbol{\gamma}}(x), b \rangle \, \boldsymbol{\gamma}(dx, dy) = 0 \quad \textit{for all } b \in \mathbf{R}^d. \tag{2.7}$$

Proof. To prove the existence of a minimizer, note that (2.4) is equivalent to

$$\|\bar{\gamma} - \mathbf{r}_{\boldsymbol{\gamma}}\|_{\mathscr{L}^2(\mathbf{R}^d, \varrho)} \leqslant \|\bar{\gamma} - \mathbf{r}\|_{\mathscr{L}^2(\mathbf{R}^d, \varrho)} \quad \text{for all } \mathbf{r} \in \mathscr{C}_\varrho, \tag{2.8}$$

where $\boldsymbol{\gamma}(dx, dy) =: \gamma_x(dy) \, \varrho(dx)$ is the disintegration of $\boldsymbol{\gamma}$ and

$$\bar{\gamma}(x) := \int_{\mathbf{R}^d} y \, \gamma_x(dy) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d$$

is the barycentric projection. This shows that the optimal transport map $\mathbf{r}_{\boldsymbol{\gamma}} \in \mathscr{C}_\varrho$ is the metric projection of the barycentric projection $\bar{\gamma} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ onto the closed, convex subset $\mathscr{C}_\varrho$ of $\mathscr{L}^2(\mathbf{R}^d, \varrho)$. Its existence and uniqueness are well-known.

Since $\mathscr{C}_\varrho$ is in fact a cone in the Hilbert space $\mathscr{L}^2(\mathbf{R}^d,\varrho)$, the minimizer $\mathbf{r}_\varrho$ is characterized by the following condition (see Lemma 1.1 in [27]):

$$\int_{\mathbf{R}^d} \langle \bar\gamma(x) - \mathbf{r}_{\boldsymbol\gamma}(x), \mathbf{r}_{\boldsymbol\gamma}(x) - \mathbf{r}(x)\rangle\, \varrho(dx) \geqslant 0 \quad \text{for all } \mathbf{r}\in\mathscr{C}_\varrho.$$

By definition of $\bar\gamma$, this inequality is equivalent to

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbf{r}_{\boldsymbol\gamma}(x), \mathbf{r}_{\boldsymbol\gamma}(x) - \mathbf{r}(x)\rangle\, \boldsymbol\gamma(dx,dy) \geqslant 0 \quad \text{for all } \mathbf{r}\in\mathscr{C}_\varrho. \tag{2.9}$$

Using this estimate for $\mathbf{r}=0$ and $\mathbf{r}=2\mathbf{r}_{\boldsymbol\gamma}$, which are both in $\mathscr{C}_\varrho$ because of its cone structure, we obtain (2.5). Inequality (2.6) follows from the choice $\mathbf{r}=\mathrm{id}$.

Notice that using (2.5) in (2.9), we find that

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbf{r}_{\boldsymbol\gamma}(x), \mathbf{r}(x)\rangle\, \boldsymbol\gamma(dx,dy) \leqslant 0 \quad \text{for all } \mathbf{r}\in\mathscr{C}_\varrho. \tag{2.10}$$

Applying this estimate with $\mathbf{r}=\pm b$ and $b\in\mathbf{R}^d$, we get (2.7) □

Remark 2.8. There is an equivalent characterization of $\mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$ as

$$\text{the closure of } \Big\{\lambda(\mathbf{r}-\mathrm{id})\colon \mathbf{r}\in\mathscr{C}_\varrho, \lambda>0\Big\} \text{ in the } \mathscr{L}^2(\mathbf{R}^d,\varrho)\text{-norm;}$$

see Theorem 8.5.1 in [5]. For all optimal transport maps $\mathbf{r}\in\mathscr{C}_\varrho$ we have that $\mathrm{id}+\mathbf{r}$ is again optimal because $\mathscr{C}_\varrho$ is a convex cone. Hence $\mathscr{C}_\varrho\subset\mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$. Using the decomposition (1.13), we then find that inequality (2.8) is equivalent to

$$\|P_\varrho(\bar\gamma) - \mathbf{r}_{\boldsymbol\gamma}\|_{\mathscr{L}^2(\mathbf{R}^d,\varrho)} \leqslant \|P_\varrho(\bar\gamma) - \mathbf{r}\|_{\mathscr{L}^2(\mathbf{R}^d,\varrho)} \quad \text{for all } \mathbf{r}\in\mathscr{C}_\varrho, \tag{2.11}$$

where $P_\varrho\colon \mathscr{L}^2(\mathbf{R}^d,\varrho)\longrightarrow\mathbb{T}_\varrho\mathscr{P}(\mathbf{R}^d)$ is the orthogonal projection. To compute the metric projection of $\boldsymbol\gamma$ onto the cone of optimal transport maps, we therefore only need the tangent component $P_\varrho(\bar\gamma)$ of the barycentric projection of $\boldsymbol\gamma$.

Remark 2.9 (Polar Cone). Inequality (2.10) is equivalent to

$$\int_{\mathbf{R}^d} \langle \bar\gamma - \mathbf{r}_{\boldsymbol\gamma}(x), \mathbf{r}(x)\rangle\, \varrho(dx) \leqslant 0 \quad \text{for all } \mathbf{r}\in\mathscr{C}_\varrho, \tag{2.12}$$

which means that the difference $\bar\gamma-\mathbf{r}_{\boldsymbol\gamma}$ is an element of the polar cone of $\mathscr{C}_\varrho$. Again we may replace $\bar\gamma$ by its tangent component $P_\varrho(\bar\gamma)$; see Remark 2.8.

We now want to rewrite the discussion above in terms of *velocities* instead of plans. To this end, consider the bijection $H_\tau\colon \mathbf{R}^{2d}\longrightarrow\mathbf{R}^{2d}$ defined as

$$H_\tau(x,\xi) := (x, x+\tau\xi) \quad \text{for all } (x,\xi)\in\mathbf{R}^{2d},$$

where $\tau>0$ is some fixed number. Its inverse function is given by

$$H_\tau^{-1}(x,y) := \left(x, \frac{y-x}{\tau}\right) \quad \text{for all } (x,y)\in\mathbf{R}^{2d},$$

and the pushforward under H_τ is an automorphism of $\mathscr{P}_\varrho(\mathbf{R}^{2d})$ for any $\varrho\in\mathscr{P}(\mathbf{R}^d)$. To every transport map $\mathbf{r}\in\mathscr{L}^2(\mathbf{R}^d,\varrho)$ we can assign a velocity $\mathbf{u}\in\mathscr{L}^2(\mathbf{R}^d,\varrho)$ by defining $\mathbf{r} =: \mathrm{id}+\tau\mathbf{u}$. In a similar fashion, to every transport plan $\boldsymbol\gamma\in\mathscr{P}_\varrho(\mathbf{R}^{2d})$ we can assign a generalized velocity by $\boldsymbol\mu := (H_\tau^{-1})_\#\boldsymbol\gamma\in\mathscr{P}_\varrho(\mathbf{R}^{2d})$.

Theorem 2.10 (Velocity Projection). *Let $\varrho \in \mathscr{P}(\mathbf{R}^d)$ and $\tau > 0$ be given. For any measure $\boldsymbol{\mu} \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ we consider the transport plan $\boldsymbol{\gamma}_{\boldsymbol{\mu},\tau} := (H_\tau)_\# \boldsymbol{\mu} \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$, and we let $\mathbf{r}_{\boldsymbol{\mu},\tau} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ be the uniquely determined optimal transport map of Theorem 2.7, with $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{\boldsymbol{\mu},\tau}$. Then we define a velocity $\mathbf{u}_{\boldsymbol{\mu},\tau} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ by*

$$\mathbf{r}_{\boldsymbol{\mu},\tau} =: \mathrm{id} + \tau \mathbf{u}_{\boldsymbol{\mu},\tau} \quad \varrho\text{-a.e.}, \tag{2.13}$$

which has the following properties:

$$\int_{\mathbf{R}^{2d}} \langle \xi - \mathbf{u}_{\boldsymbol{\mu},\tau}(x), x + \tau \mathbf{u}_{\boldsymbol{\mu},\tau}(x) \rangle \, \boldsymbol{\mu}(dx, d\xi) = 0, \tag{2.14}$$

$$\int_{\mathbf{R}^d} |\mathbf{u}_{\boldsymbol{\mu},\tau}(x)|^2 \, \varrho(dx) + \int_{\mathbf{R}^{2d}} |\xi - \mathbf{u}_{\boldsymbol{\mu},\tau}(x)|^2 \, \boldsymbol{\mu}(dx, d\xi) \leqslant \int_{\mathbf{R}^{2d}} |\xi|^2 \, \boldsymbol{\mu}(dx, d\xi), \tag{2.15}$$

$$\int_{\mathbf{R}^{2d}} \langle \xi - \mathbf{u}_{\boldsymbol{\mu},\tau}(x), b \rangle \, \boldsymbol{\mu}(dx, d\xi) = 0 \quad \text{for all } b \in \mathbf{R}^d. \tag{2.16}$$

Proof. The equalities (2.14) and (2.16) follow immediately from the definitions and Theorem 2.7. To prove (2.15), notice first that (2.6) and (2.13) imply that

$$\begin{aligned}
0 &\leqslant \int_{\mathbf{R}^{2d}} \langle y - \mathbf{r}_{\boldsymbol{\mu},\tau}(x), \mathbf{r}_{\boldsymbol{\mu},\tau}(x) - x \rangle \, \boldsymbol{\gamma}_{\boldsymbol{\mu},\tau}(dx, dy) \\
&= \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - (x + \tau \mathbf{u}_{\boldsymbol{\mu},\tau}(x)), \tau \mathbf{u}_{\boldsymbol{\mu},\tau}(x) \rangle \, \boldsymbol{\mu}(dx, d\xi) \\
&= \tau^2 \int_{\mathbf{R}^{2d}} \langle \xi - \mathbf{u}_{\boldsymbol{\mu},\tau}(x), \mathbf{u}_{\boldsymbol{\mu},\tau}(x) \rangle \, \boldsymbol{\mu}(dx, d\xi),
\end{aligned}$$

and so

$$-2 \int_{\mathbf{R}^{2d}} \langle \xi, \mathbf{u}_{\boldsymbol{\mu},\tau}(x) \rangle \, \boldsymbol{\mu}(dx, d\xi) \leqslant -2 \int_{\mathbf{R}^d} |\mathbf{u}_{\boldsymbol{\mu},\tau}(x)|^2 \, \varrho(dx). \tag{2.17}$$

Expanding the quadratic term in (2.15) and using (2.17), we obtain the result. □

3. Stress Tensor

In this section, we will show that the difference in Theorem 2.10 between the velocity distribution determined by the measure $\boldsymbol{\mu} \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ and the optimal transport velocity $\mathbf{u}_{\boldsymbol{\mu},\tau} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ can be decomposed into the divergence of a matrix-valued measure field, which we will call a stress tensor field, and a divergence-free component. We will derive a precise estimate on the size of the stress tensor.

We write $\mathcal{S}^d$ for the set of all symmetric $(d \times d)$-matrices, and $\mathcal{S}^d_+$ resp. $\mathcal{S}^d_{++}$ for the subsets of positive semi-definite resp. strictly positive definite matrices. For any $A, B \in \mathcal{S}^d$ we denote by $\langle A, B \rangle := \mathrm{tr}(AB)$ the inner product.

Theorem 3.1 (Stress Tensor). *Fix $\tau > 0$. For any $\varrho \in \mathscr{P}(\mathbf{R}^d)$ and $\boldsymbol{\mu} \in \mathscr{P}_\varrho(\mathbf{R}^{2d})$ let $\mathbf{u}_{\boldsymbol{\mu},\tau} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ be the velocity projection of Theorem 2.10. Let*

$$\boldsymbol{\zeta}_{\boldsymbol{\mu},\tau} := \bar{\mu} - \mathbf{u}_{\boldsymbol{\mu},\tau} \in \mathscr{L}^2(\mathbf{R}^d, \varrho),$$

where $\boldsymbol{\mu}(dx, d\xi) =: \mu_x(d\xi) \, \varrho(dx)$ is the disintegration of $\boldsymbol{\mu}$ and

$$\bar{\mu}(x) := \int_{\mathbf{R}^d} \xi \, \mu_x(d\xi) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d.$$

Then there exists a measure $\boldsymbol{\sigma}_{\boldsymbol{\mu},\tau} \in \mathscr{M}(\mathbf{R}^d, \mathcal{S}_+^d)$ with the property that

$$\langle D^2, \boldsymbol{\sigma}_{\boldsymbol{\mu},\tau}\rangle = \nabla \cdot (\varrho \boldsymbol{\zeta}_{\boldsymbol{\mu},\tau}) \quad \textit{in } \mathscr{D}'(\mathbf{R}^d), \tag{3.1}$$

$$\int_{\mathbf{R}^d} \operatorname{tr}(\boldsymbol{\sigma}_{\boldsymbol{\mu},\tau}) \leqslant -\int_{\mathbf{R}^d} \langle \mathrm{id}, \boldsymbol{\zeta}_{\boldsymbol{\mu},\tau}\rangle \varrho \, dx. \tag{3.2}$$

Using (2.14) *in* (3.2)*, and then inequality* (2.15)*, we obtain*

$$\begin{aligned}\int_{\mathbf{R}^d} \operatorname{tr}(\boldsymbol{\sigma}_{\boldsymbol{\mu},\tau}) \leqslant \tau \bigg(\int_{\mathbf{R}^{2d}} |\xi|^2 \, \boldsymbol{\mu}(dx, d\xi) \bigg)^{1/2} \\ \times \bigg(\int_{\mathbf{R}^{2d}} |\xi|^2 \, \boldsymbol{\mu}(dx, d\xi) - \int_{\mathbf{R}^d} |\mathbf{u}_{\boldsymbol{\mu},\tau}|^2 \, \varrho \bigg)^{1/2}\end{aligned} \tag{3.3}$$

Proof. We modify the argument in [17]. For any measurable function $f \colon \mathbf{R}^d \longrightarrow \mathbf{R}$ we denote by $[f] \in \mathscr{D}'(\mathbf{R}^d)$ the distribution defined by integration against f:

$$[f](\varphi) := \int_{\mathbf{R}^d} f \varphi \, dx \quad \text{for all } \varphi \in \mathscr{D}(\mathbf{R}^d).$$

To simplify notation, we will omit the subscript $\boldsymbol{\mu}, \tau$ in the following.

Step 1. It is well-known [11] that a distribution $T \in \mathscr{D}'(\mathbf{R}^d)$ is induced by a convex function if and only if $D^2 T \in \mathscr{D}'(\mathbf{R}^d, \mathcal{S}_+^d)$. That is, if $T(D^2\varphi)$ is a symmetric and positive semi-definite matrix for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$ with $\varphi \geqslant 0$, so that

$$\langle v, T(D^2\varphi) v\rangle \geqslant 0 \quad \text{for all } v \in \mathbf{R}^d. \tag{3.4}$$

This is equivalent to the following statement: the distribution T is induced by a convex function if and only if for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$ with $\varphi \geqslant 0$, we have

$$\langle A, T(D^2\varphi)\rangle \geqslant 0 \quad \text{for all } A \in \mathcal{S}_{++}^d \text{ with } \operatorname{tr}(A) = d. \tag{3.5}$$

Indeed, assume that (3.4) holds. Any $A \in \mathcal{S}_{++}^d$ can be written in the form

$$A = \sum_{i=1}^d \lambda_i \, e_i \otimes e_i, \tag{3.6}$$

where $\lambda_i > 0$ are the eigenvalues of A and e_i are the corresponding eigenvectors, which form an orthonormal basis for $\mathbf{R}^d$. This implies that

$$\langle A, T(D^2\varphi)\rangle = \sum_{i=1}^d \lambda_i \, \langle e_i, T(D^2\varphi) e_i\rangle,$$

which is nonnegative because of (3.4), and so (3.5) follows.

Conversely, assume that (3.4) does not hold and there exist a $\varphi \in \mathscr{D}(\mathbf{R}^d)$ with $\varphi \geqslant 0$ and a vector $v \in \mathbf{R}^d$ such that $\langle v, T(D^2\varphi) v\rangle =: \delta < 0$. Without loss of generality, we may assume that $|v| = 1$. Then we define

$$\alpha := \sup \Big\{ |\langle w, T(D^2\varphi) w\rangle| \colon w \in \mathbf{R}^d, |w| = 1 \Big\},$$

which is a finite number. Now fix some orthonormal basis $\{e_1 := v, e_2, \dots, e_n\}$ and an $\varepsilon > 0$ sufficiently small such that $\delta + (n-1)\alpha\varepsilon < 0$. Let

$$A := \frac{d}{1 + (n-1)\varepsilon} \bigg(v \otimes v + \varepsilon \sum_{i=2}^n e_i \otimes e_i \bigg).$$

Then $A \in \mathcal{S}^d_{++}$ with $\operatorname{tr}(A) = d$ and we have that

$$\begin{aligned}\langle A, T(D^2\varphi)\rangle &= \frac{d}{1+(n-1)\varepsilon}\Bigg(\langle v, T(D^2\varphi)v\rangle + \varepsilon\sum_{i=2}^n \langle e_i, T(D^2\varphi)e_i\rangle\Bigg)\\ &\leqslant d\frac{\delta+(n-1)\alpha\varepsilon}{1+(n-1)\varepsilon} < 0.\end{aligned}$$

Therefore, if (3.4) does not hold, then (3.5) does not hold either.

This proves the following identity:

$$\mathscr{D}'_{\mathrm{conv}}(\mathbf{R}^d) = \bigcap\Big\{\mathscr{D}'_A(\mathbf{R}^d)\colon A \in \mathcal{S}^d_{++} \text{ with } \operatorname{tr}(A) = d\Big\}, \tag{3.7}$$

where

$$\mathscr{D}'_{\mathrm{conv}}(\mathbf{R}^d) := \Big\{T \in \mathscr{D}'(\mathbf{R}^d)\colon T = [\phi] \text{ with } \phi\colon \mathbf{R}^d \longrightarrow \mathbf{R} \text{ u.s.c. convex}\Big\}, \tag{3.8}$$

$$\mathscr{D}'_A(\mathbf{R}^d) := \Big\{T \in \mathscr{D}'(\mathbf{R}^d)\colon \langle A, D^2\rangle T \geqslant 0 \text{ in } \mathscr{D}'(\mathbf{R}^d)\Big\}. \tag{3.9}$$

Here we assumed without loss of generality that the convex functions in (3.8) are upper semicontinuous. Recall that a convex function is continuous in the interior of its domain. By linearity, we have $\langle A, T(D^2\varphi)\rangle = T(\langle A, D^2\varphi\rangle)$ for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$. It is easy to check that both (3.8) and (3.9) are convex cones.

Step 2. Any $A \in \mathcal{S}^d_{++}$ admits a unique square root given by

$$B := \sum_{i=1}^d \sqrt{\lambda_i}\, e_i \otimes e_i \in \mathcal{S}^d_{++},$$

where λ_i, e_i are the eigenvalues and eigenvectors of A introduced in (3.6). In order to simplify notation, we will work with B instead of $A = B^2$ in the following.

Defining the linear map $\ell_B(x) := Bx$ for all $x \in \mathbf{R}^d$, we obtain that

$$\langle B^2, D^2\rangle(\varphi\circ\ell_{B^{-1}}) = (\Delta\varphi)\circ\ell_{B^{-1}} \quad \text{for all } \varphi \in \mathscr{D}(\mathbf{R}^d).$$

Consider now any $T \in \mathscr{D}'_{B^2}(\mathbf{R}^d)$. Then

$$\begin{aligned}\Big(\Delta(T\circ\ell_B)\Big)(\varphi) &= T\bigg(\frac{(\Delta\varphi)\circ\ell_{B^{-1}}}{|\det B|}\bigg)\\ &= T\bigg(\frac{\langle B^2, D^2\rangle(\varphi\circ\ell_{B^{-1}})}{|\det B|}\bigg) = \Big(\langle B^2, D^2\rangle T\Big)\bigg(\frac{\varphi\circ\ell_{B^{-1}}}{|\det B|}\bigg).\end{aligned} \tag{3.10}$$

Notice that $\varphi \in \mathscr{D}(\mathbf{R}^d)$ implies $\varphi_B := \varphi\circ\ell_{B^{-1}}/|\det B| \in \mathscr{D}(\mathbf{R}^d)$, and that $\varphi_B \geqslant 0$ whenever $\varphi \geqslant 0$. Therefore the distribution $T_B := T\circ\ell_B \in \mathscr{D}'(\mathbf{R}^d)$ is *subharmonic*. It is well-known [10] that any subharmonic distribution is induced by a subharmonic function, so that $T_B = [\phi_B]$, where $\phi_B\colon \mathbf{R}^d \longrightarrow \mathbf{R}$ may be assumed u.s.c. Since

$$T\bigg(\frac{\varphi\circ\ell_{B^{-1}}}{|\det B|}\bigg) = T_B(\varphi) = \int_{\mathbf{R}^d}\phi_B\varphi\,dx = \int_{\mathbf{R}^d}\big(\phi_B\circ\ell_{B^{-1}}\big)\bigg(\frac{\varphi\circ\ell_{B^{-1}}}{|\det B|}\bigg)\,dx.$$

for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$, we conclude that $T = [\phi]$ with $\phi := \phi_B\circ\ell_{B^{-1}}$.

Step 3. Assume now that ϱ is smooth and $\varrho(x) > 0$ for all $x \in \mathbf{R}^d$. Let

$$\mathscr{C}_{\varrho,B} := \Big\{\mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d,\varrho)\colon [\mathbf{r}] = \nabla T \text{ with } T \in \mathscr{D}'_{B^2}(\mathbf{R}^d)\Big\}$$

for all $B \in \mathcal{S}^d_{++}$ with $\operatorname{tr}(B^2) = d$. Then $\mathscr{C}_{\varrho,B}$ is a convex cone. To prove that it is closed in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$, consider a sequence $\mathbf{r}^k \in \mathscr{C}_{\varrho,B}$ with $[\mathbf{r}^k] =: \nabla T^k$ and

$$\mathbf{r}^k \longrightarrow \mathbf{r} \quad \text{in } \mathscr{L}^2(\mathbf{R}^d, \varrho)$$

for some $\mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$. Since ϱ is smooth and positive, it is uniformly positive on compact sets, and so convergence in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$ implies convergence in $\mathscr{L}^2_{\mathrm{loc}}(\mathbf{R}^d)$, which in turn implies convergence in the distributional sense. Since

$$\partial_i[\mathbf{r}^k]_j - \partial_j[\mathbf{r}^k]_i = \partial_i(\partial_j T^k) - \partial_j(\partial_i T^k) = 0 \quad \text{in } \mathscr{D}'(\mathbf{R}^d)$$

for all $1 \leqslant i, j \leqslant d$ and all k, we obtain that

$$\partial_i[\mathbf{r}]_j - \partial_j[\mathbf{r}]_i = 0 \quad \text{in } \mathscr{D}'(\mathbf{R}^d).$$

Therefore there is a distribution $T \in \mathscr{D}'(\mathbf{R}^d)$ such that $[\mathbf{r}] = \nabla T$; see [24]. Since

$$\begin{array}{ccc} \Big(\langle B^2, D^2\rangle T^k\Big)(\varphi) & & \Big(\langle B^2, D^2\rangle T\Big)(\varphi) \\ \Big\| & & \Big\| \\ -\displaystyle\int_{\mathbf{R}^d} \langle \mathbf{r}^k, B^2\nabla\varphi\rangle\, dx & \longrightarrow & -\displaystyle\int_{\mathbf{R}^d} \langle \mathbf{r}, B^2\nabla\varphi\rangle\, dx \end{array}$$

for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$ (in particular, those with $\varphi \geqslant 0$), we have that $T \in \mathscr{D}'_{B^2}(\mathbf{R}^d)$. Therefore $\mathscr{C}_{\varrho,B}$ is a closed, convex cone for all $B \in \mathcal{S}^d_{++}$ with $\operatorname{tr}(B^2) = d$.

Let $\mathscr{C}_\varrho$ be the closed, convex cone of optimal transport maps introduced in (2.1). Since ϱ is absolutely continuous with respect to the Lebesgue measure, every $\mathbf{r} \in \mathscr{C}_\varrho$ is defined only almost everywhere with respect to $\mathcal{L}^d$. On the other hand, every cyclically monotone subset in $\mathbf{R}^d \times \mathbf{R}^d$ is contained in the *subdifferential* of a convex function, which is single-valued everywhere outside a $\mathcal{H}^{d-1}$-rectifiable set; see [1]. An optimal transport map in $\mathscr{C}_\varrho$ with ϱ absolutely continuous with respect to $\mathcal{L}^d$ is therefore uniquely determined as the gradient of a convex function. Hence

$$\mathscr{C}_\varrho = \Big\{ \mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d, \varrho) \colon [\mathbf{r}] = \nabla T \text{ with } T \in \mathscr{D}'_{\mathrm{conv}}(\mathbf{R}^d) \Big\}.$$

As a consequence of (3.7), we then find that

$$\mathscr{C}_\varrho = \bigcap \Big\{ \mathscr{C}_{\varrho,B} \colon B \in \mathcal{S}^d_{++} \text{ with } \operatorname{tr}(B^2) = d \Big\}. \tag{3.11}$$

To simplify the notation, in the following we will no longer distinguish between a locally integrable function and the distribution it induces by integration.

Step 4. We now introduce the polar cones

$$\mathscr{C}_\varrho^\perp := \Big\{ \boldsymbol{\zeta} \in \mathscr{L}^2(\mathbf{R}^d, \varrho) \colon \int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}, \mathbf{r}\rangle \varrho \leqslant 0 \text{ for all } \mathbf{r} \in \mathscr{C}_\varrho \Big\},$$
$$\mathscr{C}_{\varrho,B}^\perp := \Big\{ \boldsymbol{\zeta} \in \mathscr{L}^2(\mathbf{R}^d, \varrho) \colon \int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}, \mathbf{r}\rangle \varrho \leqslant 0 \text{ for all } \mathbf{r} \in \mathscr{C}_{\varrho,B} \Big\}$$

for all $B \in \mathcal{S}^d_{++}$ with $\operatorname{tr}(B^2) = d$. Then (3.11) implies that

$$\mathscr{C}_\varrho^\perp = \overline{\operatorname{conv} \bigcup \Big\{ \mathscr{C}_{\varrho,B}^\perp \colon B \in \mathcal{S}^d_{++} \text{ with } \operatorname{tr}(B^2) = d \Big\}}^{\mathscr{L}^2(\mathbf{R}^d, \varrho)} \tag{3.12}$$

(that is, the polar cone of $\mathscr{C}_\varrho$ is the closure in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$ of the convex hull of the polar cones of $\mathscr{C}_{\varrho,B}$); see Corollary 2 to Lemma 2.1 in [27] for a proof.

We now use the following characterization:

Lemma 3.2. *For all $B \in \mathcal{S}^d_{++}$ such that $\operatorname{tr}(B^2) = d$ and for all $\boldsymbol{\zeta} \in \mathscr{C}^\perp_{\varrho,B}$ there exists a nonnegative function $f \in \mathscr{L}^1(\mathbf{R}^d)$ with the property that*

$$\langle B^2, D^2\rangle f = \nabla\cdot(\varrho\boldsymbol{\zeta}) \quad \text{in } \mathscr{D}'(\mathbf{R}^d), \tag{3.13}$$

$$d\int_{\mathbf{R}^d} f(x)\,dx = -\int_{\mathbf{R}^d}\langle x, \boldsymbol{\zeta}(x)\rangle\varrho(x)\,dx. \tag{3.14}$$

Before proving Lemma 3.2 let us finish the proof of the proposition. Let $\boldsymbol{\zeta} \in \mathscr{C}^\perp_\varrho$ be given. By (3.12), there then exist sequences of integers $N_k \in \mathbf{N}$ and

$$B_{k,i} \in \mathcal{S}^d_{++} \text{ with } \operatorname{tr}(B^2_{k,i}) = d, \quad \lambda_{k,i} \in [0,1], \quad \boldsymbol{\zeta}_{k,i} \in \mathscr{C}^\perp_{\varrho,B_{k,i}}$$

for all $1 \leqslant i \leqslant N_k$, with the property that $\sum_{i=1}^{N_k}\lambda_{k,i} = 1$ and

$$\boldsymbol{\zeta}_k := \sum_{i=1}^{N_k}\lambda_{k,i}\boldsymbol{\zeta}_{k,i} \longrightarrow \boldsymbol{\zeta} \quad \text{in } \mathscr{L}^2(\mathbf{R}^d,\varrho). \tag{3.15}$$

By Lemma 3.2, for every $\boldsymbol{\zeta}_{k,i}$ there exists a nonnegative $f_{k,i} \in \mathscr{L}^1(\mathbf{R}^d)$ and

$$\langle D^2, f_{k,i}B^2_{k,i}\rangle = \nabla\cdot(\varrho\boldsymbol{\zeta}_{k,i}) \quad \text{in } \mathscr{D}'(\mathbf{R}^d).$$

$$d\int_{\mathbf{R}^d} f_{k,i}(x)\,dx = -\int_{\mathbf{R}^d}\langle x, \boldsymbol{\zeta}_{k.i}(x)\rangle\varrho(x)\,dx.$$

We now define

$$\boldsymbol{\sigma}_k := \sum_{i=1}^{N_k}\lambda_{k,i}f_{k,i}B^2_{k,i} \in \mathscr{L}^1(\mathbf{R}^d,\mathcal{S}^d_+).$$

By linearity, we then have

$$\langle D^2, \boldsymbol{\sigma}_k\rangle = \nabla\cdot(\varrho\boldsymbol{\zeta}_k) \quad \text{in } \mathscr{D}'(\mathbf{R}^d), \tag{3.16}$$

$$\int_{\mathbf{R}^d}\operatorname{tr}(\boldsymbol{\sigma}_k(x))\,dx = -\int_{\mathbf{R}^d}\langle x, \boldsymbol{\zeta}_k(x)\rangle\varrho(x)\,dx. \tag{3.17}$$

Because of (3.15), the right-hand sides of (3.16) and (3.17) converge to the right-hand sides of (3.1) and (3.2). In particular, the $\mathscr{L}^1(\mathbf{R}^d)$-norm of $\operatorname{tr}(\boldsymbol{\sigma}_k)$ is uniformly bounded. Recall that for positive definite matrices the trace is equivalent to the trace-norm and thus to any matrix norm. Therefore the sequence $\{\boldsymbol{\sigma}_k\}$ is uniformly bounded in $\mathscr{M}(\mathbf{R}^d,\mathcal{S}^d_+)$. By Banach-Alaoglu, there exists a subsequence (which we still denote by $\{\boldsymbol{\sigma}_k\}$ for simplicity) and a measure $\boldsymbol{\sigma} \in \mathscr{M}(\mathbf{R}^d,\mathcal{S}^d_+)$ such that

$$\boldsymbol{\sigma}_k \rightharpoonup \boldsymbol{\sigma} \quad \text{weak* in } \mathscr{M}(\mathbf{R}^d,\mathcal{S}^d_+). \tag{3.18}$$

This implies that $\langle D^2,\boldsymbol{\sigma}_k\rangle \rightharpoonup \langle D^2,\boldsymbol{\sigma}\rangle$ in $\mathscr{D}'(\mathbf{R}^d)$, and so (3.1) follows.

To prove the estimate (3.2), we first note that

$$\int_{\mathbf{R}^d}\operatorname{tr}(\boldsymbol{\sigma}) = \sup\bigg\{\int_{\mathbf{R}^d}\langle\omega,\boldsymbol{\sigma}\rangle\colon \omega \in \mathscr{C}_{\mathrm{b}}(\mathbf{R}^d,\mathcal{S}^d),\ \rho(\omega(x)) \leqslant 1 \text{ for all } x \in \mathbf{R}^d\bigg\},$$

where $\rho(A) := \max\{\langle v, Av\rangle\colon |v| \leqslant 1\}$ for all $A \in \mathcal{S}^d$; see [7]. Since $\mathscr{C}_{\mathrm{b}}(\mathbf{R}^d,\mathcal{S}^d)$ is separable, the functional $\boldsymbol{\sigma} \mapsto \int_{\mathbf{R}^d}\operatorname{tr}(\boldsymbol{\sigma})$ is lower semicontinuous with respect to the weak* convergence of measures. Because of (3.18), we therefore obtain

$$\int_{\mathbf{R}^d}\operatorname{tr}(\boldsymbol{\sigma}) \leqslant \liminf_{k\to\infty}\int_{\mathbf{R}^d}\operatorname{tr}(\boldsymbol{\sigma}_k).$$

Step 5. We consider now a general density $\varrho \in \mathscr{P}(\mathbf{R}^d)$ and $\boldsymbol{\zeta} \in \mathscr{C}_\varrho^\perp$. Let

$$\varphi_\varepsilon(x) := \varepsilon^{-d}\varphi_1(x/\varepsilon) \quad \text{for all } x \in \mathbf{R}^d \text{ and } \varepsilon > 0,$$

where $\varphi_1(x) := \omega_d \exp(-|x|^2)$ and ω_d is chosen such that $\int_{\mathbf{R}^d} \varphi_1(x)\, dx = 1$. Let

$$\varrho_\varepsilon := \varrho \star \varphi_\varepsilon \quad \text{and} \quad (\varrho\boldsymbol{\zeta})_\varepsilon := (\varrho\boldsymbol{\zeta}) \star \varphi_\varepsilon.$$

Since $\varrho_\varepsilon(x) > 0$ for all $x \in \mathbf{R}^d$, we can define $\boldsymbol{\zeta}_\varepsilon := (\varrho\boldsymbol{\zeta})_\varepsilon/\varrho_\varepsilon$, which is smooth for all $\varepsilon > 0$. Since φ_1 is an even function, we have that

$$\int_{\mathbf{R}^d} |x|^2 \varrho_\varepsilon(x)\, dx \leqslant \varepsilon^2 \int_{\mathbf{R}^d} |z|^2 \varphi_1(z)\, dz + \int_{\mathbf{R}^d} |y|^2\, \varrho(dy),$$

and so $\varrho_\varepsilon \in \mathscr{P}(\mathbf{R}^d)$ for all $\varepsilon > 0$. We now use the following identity:

$$\int_{\mathbf{R}^d} |\boldsymbol{\zeta}|^2 \varrho = \sup_\xi \left\{ \int_{\mathbf{R}^d} \Big(-|\xi|^2 \varrho + 2\langle \varrho\boldsymbol{\zeta}, \xi\rangle \Big) \right\}, \tag{3.19}$$

which holds for all densities $\varrho \in \mathscr{P}(\mathbf{R}^d)$ and vector fields $\boldsymbol{\zeta} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$; see [6]. The sup in (3.19) is taken over all vector-valued $\xi \in \mathscr{C}_{\mathrm{b}}(\mathbf{R}^d)$. This yields

$$\begin{aligned}\int_{\mathbf{R}^d} |\boldsymbol{\zeta}_\varepsilon|^2 \varrho_\varepsilon\, dx &= \sup_\xi \left\{ \int_{\mathbf{R}^d} \Big(-\big(|\xi|^2\big)_\varepsilon + 2\langle \boldsymbol{\zeta}, \xi_\varepsilon\rangle \Big) \varrho \right\} \\ &\leqslant \sup_\xi \left\{ \int_{\mathbf{R}^d} \Big(-|\xi_\varepsilon|^2 + 2\langle \boldsymbol{\zeta}, \xi_\varepsilon\rangle \Big) \varrho \right\} \leqslant \int_{\mathbf{R}^d} |\boldsymbol{\zeta}|^2 \varrho,\end{aligned}$$

where $\xi_\varepsilon := \xi \star \varphi_\varepsilon$. We used Jensen's ineqality to estimate

$$\begin{aligned}\big(|\xi|^2\big)_\varepsilon(x) := \int_{\mathbf{R}^d} |\xi(y)|^2 \varphi_\varepsilon(x-y)\, dy &\geqslant \bigg(\int_{\mathbf{R}^d} |\xi(y)| \varphi_\varepsilon(x-y)\, dy \bigg)^2 \\ &\geqslant \bigg| \int_{\mathbf{R}^d} \xi(y) \varphi_\varepsilon(x-y)\, dy \bigg|^2 = |\xi_\varepsilon(x)|^2\end{aligned}$$

for all $x \in \mathbf{R}^d$. Hence $\boldsymbol{\zeta}_\varepsilon \in \mathscr{L}^2(\mathbf{R}^d, \varrho_\varepsilon)$.

Consider now an $\mathbf{r} \in \mathscr{C}_{\varrho_\varepsilon}$. Note that since ϱ_ε is smooth, every optimal transport plan over ϱ_ε is in fact induced by a map. To prove that the convolution $\mathbf{r}_\varepsilon := \mathbf{r} \star \varphi_\varepsilon$ is well-defined, let f be any component of the vector-valued function $\mathbf{r}$. We denote by $f_+ := \max\{f, 0\}$ its positive part. We choose a sequence $\{s^k\}$ of simple functions with $0 \leqslant s^k(x) \leqslant s^{k+1}(x) \leqslant f_+(x)$ for a.e. $x \in \mathbf{R}^d$ and $k \in \mathbf{N}$, and $s^k(x) \to f_+(x)$ as $k \to \infty$. Since s^k is bounded, the convolution $s^k_\varepsilon := s^k \star \varphi_\varepsilon$ is well-defined. By Jensen's inequality, we have that $s^k_\varepsilon(x)^2 \leqslant (s^k)^2 \star \varphi_\varepsilon(x)$ and therefore

$$\int_{\mathbf{R}^d} s^k_\varepsilon(x)^2\, \varrho(dx) \leqslant \int_{\mathbf{R}^d} s^k(x)^2 \varrho_\varepsilon(x)\, dx \leqslant \int_{\mathbf{R}^d} f_+(x)^2 \varrho_\varepsilon(x)\, dx, \tag{3.20}$$

which is finite since $\mathbf{r} \in \mathscr{L}^2(\mathbf{R}^d, \varrho_\varepsilon)$. By monotone convergence, we have that

$$s^k_\varepsilon(x) = \int_{\mathbf{R}^d} s^k(y) \varphi_\varepsilon(x-y)\, dy \longrightarrow \int_{\mathbf{R}^d} f_+(y) \varphi_\varepsilon(x-y)\, dy =: f_{+,\varepsilon}(x)$$

for all $x \in \mathbf{R}^d$. By Fatou's lemma and (3.20), we obtain that

$$\int_{\mathbf{R}^d} f_{+,\varepsilon}(x)^2\, \varrho(dx) \leqslant \int_{\mathbf{R}^d} f_+(x)^2 \varrho_\varepsilon(x)\, dx.$$

In particular, we have that $f_{+,\varepsilon}$ is finite ϱ-a.e., and so $f_{+,\varepsilon} = f_+ \star \varphi_\varepsilon$ is well-defined. Repeating the same argument for the negative part of f and for all other components of $\mathbf{r}$ we get that $\mathbf{r}_\varepsilon := \mathbf{r} \star \varphi_\varepsilon$ is well-defined and $\mathbf{r}_\varepsilon \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$. Moreover, since

$\mathbf{r}$ coincides with the gradient of a convex function, and since convolution with a nonnegative function preserves the convexity, we conclude that $\mathbf{r}_\varepsilon$ is still an optimal transport map and thus $\mathbf{r}_\varepsilon \in \mathscr{C}_\varrho$. For all $\mathbf{r} \in \mathscr{C}_{\varrho_\varepsilon}$ we have that

$$\int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}_\varepsilon, \mathbf{r} \rangle \varrho_\varepsilon \, dx = \int_{\mathbf{R}^d} \langle (\varrho\boldsymbol{\zeta})_\varepsilon, \mathbf{r} \rangle \, dx = \int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}, \mathbf{r}_\varepsilon \rangle \, \varrho. \tag{3.21}$$

By choice of $\boldsymbol{\zeta}$, the integral in (3.21) must then be nonpositive, thus $\boldsymbol{\zeta}_\varepsilon \in \mathscr{C}_{\varrho_\varepsilon}^\perp$.

By Step 4, there exists a stress tensor field $\boldsymbol{\sigma}_\varepsilon \in \mathscr{L}^1(\mathbf{R}^d, \mathcal{S}_+^d)$ such that

$$\langle D^2, \boldsymbol{\sigma}_\varepsilon \rangle = \nabla \cdot (\varrho\boldsymbol{\zeta})_\varepsilon \quad \text{in } \mathscr{D}'(\mathbf{R}^d). \tag{3.22}$$

Moreover, we have the estimate

$$\begin{aligned} \int_{\mathbf{R}^d} \operatorname{tr}(\boldsymbol{\sigma}_\varepsilon) = -\int_{\mathbf{R}^d} \langle \mathrm{id}, (\varrho\boldsymbol{\zeta})_\varepsilon \rangle \, dx = -\int_{\mathbf{R}^d} \langle \mathrm{id}_\varepsilon, \boldsymbol{\zeta} \rangle \, \varrho \\ \leqslant \|\mathrm{id}\|_{\mathscr{L}^2(\mathbf{R}^d,\varrho)} \|\boldsymbol{\zeta}\|_{\mathscr{L}^2(\mathbf{R}^d,\varrho)}, \end{aligned} \tag{3.23}$$

where we used that $\mathrm{id}_\varepsilon := \mathrm{id} \star \varphi_\varepsilon = \mathrm{id}$ for all $\varepsilon > 0$ because φ_ε is an even function. Therefore the sequence $\{\boldsymbol{\sigma}_\varepsilon\}$ is uniformly bounded in $\mathscr{M}(\mathbf{R}^d, \mathcal{S}_+^d)$. By the Banach-Alaoglu theorem, there exists a subsequence (which we still denote by $\{\boldsymbol{\sigma}_\varepsilon\}$ for simplicity) and a measure $\boldsymbol{\sigma} \in \mathscr{M}(\mathbf{R}^d, \mathcal{S}_+^d)$ such that

$$\boldsymbol{\sigma}_\varepsilon \rightharpoonup \boldsymbol{\sigma} \quad \text{weak* in } \mathscr{M}(\mathbf{R}^d, \mathcal{S}_+^d) \text{ as } \varepsilon \to 0. \tag{3.24}$$

This implies that $\langle D^2, \boldsymbol{\sigma}_\varepsilon \rangle \rightharpoonup \langle D^2, \boldsymbol{\sigma} \rangle$ in $\mathscr{D}'(\mathbf{R}^d)$.

Since $\phi_\varepsilon := \phi \star \varphi_\varepsilon \longrightarrow \phi$ in $\mathscr{D}(\mathbf{R}^d)$, we also have

$$\int_{\mathbf{R}^d} \langle \nabla\phi, (\varrho\boldsymbol{\zeta})_\varepsilon \rangle \, dx = \int_{\mathbf{R}^d} \langle \nabla\phi_\varepsilon, \boldsymbol{\zeta} \rangle \, \varrho \longrightarrow \int_{\mathbf{R}^d} \langle \nabla\phi, \boldsymbol{\zeta} \rangle \, \varrho.$$

Identity (3.1) now follows from (3.22). To prove the estimate (3.2), we proceed as we did in Step 2, using the lower semicontinuity of the map $\boldsymbol{\sigma} \mapsto \int_{\mathbf{R}^d} \operatorname{tr}(\boldsymbol{\sigma})$ with respect to the weak* convergence of measures, and the convergence (3.24). □

Proof of Lemma 3.2. We split the proof into several steps.

Step 1. We consider first the case when $B = \mathrm{id}$, for which we have $\langle B^2, D^2 \rangle = \Delta$. Let $\boldsymbol{\zeta} \in \mathscr{C}_{\varrho,\mathrm{id}}^\perp$ be given and recall that $\varrho\boldsymbol{\zeta} \in \mathscr{L}^1(\mathbf{R}^d)$ because of Hölder's inequality. For $p - 1 > 0$ small and $r > d(1 - 1/p)$ there exists a constant $C > 0$ such that

$$\|\varrho\boldsymbol{\zeta}\|_{\mathscr{W}^{-r,p}(\mathbf{R}^d)} \leqslant C \|\varrho\boldsymbol{\zeta}\|_{\mathscr{L}^1(\mathbf{R}^d)},$$

by Sobolev embedding. If $d \geqslant 2$, then we obtain a solution of the equation

$$\Delta f = \nabla \cdot (\varrho\boldsymbol{\zeta}) \quad \text{in } \mathscr{D}'(\mathbf{R}^d) \tag{3.25}$$

by computing the integral transform

$$f := (-\Delta)^{-1/2} R \cdot (-\varrho\boldsymbol{\zeta}),$$

where $R := (-\Delta)^{-1/2}\nabla$ is the Riesz transform and $(-\Delta)^{-1/2}$ denotes the fractional integration operator. For $d = 1$ we use the Hilbert instead of the Riesz transform. Since $p > 1$ the Riesz/Hilbert transform is continuous from $\mathscr{W}^{-r,p}(\mathbf{R}^d)$ into itself. The fractional integration is continuous from $\mathscr{W}^{-r,p}(\mathbf{R}^d)$ to $\mathscr{L}^q(\mathbf{R}^d)$ with

$$q := \frac{dp}{d - p(1 - r)} < \frac{d}{d - 1}.$$

We refer the reader to [25] for further information. By choosing r sufficiently close to $d(1-1/p)$ we can obtain q arbitrarily close to $d/(d-1)$, thus q' close to d.

We claim that the function f defined above is nonnegative. To prove the claim, consider any $g \in \mathscr{D}(\mathbf{R}^d)$ with $g \geqslant 0$. If ϕ is the solution of $\Delta\phi = g$ given by

$$\phi(x) = \omega_d \int_{\mathbf{R}^d} \frac{g(y)}{|x-y|^{d-2}}\, dy \quad \text{for all } x \in \mathbf{R}^d$$

(with the usual modification for $d = 2$), then ϕ is smooth. Here ω_d is some constant depending only on d. If $d \geqslant 3$, then ϕ decays like $|x|^{2-d}$ as $|x| \to \infty$. If $d = 2$, then ϕ grows logarithmically. If $d = 1$, then ϕ is linear outside a compact set.

We now choose $\varphi \in \mathscr{D}(\mathbf{R})$ such that φ is even,

$$\varphi(s) \in [0,1] \text{ for all } s \in \mathbf{R}, \quad \text{and} \quad \varphi(s) = \begin{cases} 1 & \text{if } |s| \leqslant 1, \\ 0 & \text{if } |s| \geqslant 2. \end{cases} \tag{3.26}$$

We define $\psi_s \in \mathscr{D}(\mathbf{R}^d)$ for all $s > 0$ by

$$\psi_s(x) := \varphi(s|x|)\phi(x) \quad \text{for all } x \in \mathbf{R}^d,$$

and then test (3.25) against ψ_s. For the right-hand side, we obtain

$$\begin{aligned} -\int_{\mathbf{R}^d} \langle \nabla\psi_s(x), \boldsymbol{\zeta}(x)\rangle \varrho(x)\, dx = &- s\int_{\mathbf{R}^d} \varphi'(s|x|)\phi(x) \left\langle \frac{x}{|x|}, \boldsymbol{\zeta}(x) \right\rangle \varrho(x)\, dx \\ &- \int_{\mathbf{R}^d} \varphi(s|x|)\langle \nabla\phi(x), \boldsymbol{\zeta}(x)\rangle \varrho(x)\, dx. \end{aligned}$$

If $d \geqslant 3$, then we can estimate the first integral on the right-hand side as

$$\begin{aligned} &\left| s\int_{\mathbf{R}^d} \varphi'(s|x|)\phi(x) \left\langle \frac{x}{|x|}, \boldsymbol{\zeta}(x) \right\rangle \varrho(x)\, dx \right| \\ &\qquad \leqslant s\|\varphi'\|_{\mathscr{L}^\infty(\mathbf{R})} \|\phi\|_{\mathscr{L}^\infty(\mathbf{R}^d)} \|\varrho\boldsymbol{\zeta}\|_{\mathscr{L}^1(\mathbf{R}^d)} \longrightarrow 0 \quad \text{as } s \to 0. \end{aligned}$$

A similar estimate holds if $d \leqslant 2$ since ϕ grows at most linearly and $\boldsymbol{\zeta} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$. Moreover, by dominated convergence we obtain that

$$-\int_{\mathbf{R}^d} \varphi(s|x|)\langle \nabla\phi(x), \boldsymbol{\zeta}(x)\rangle \varrho(x)\, dx \longrightarrow -\int_{\mathbf{R}^d} \langle \nabla\phi(x), \boldsymbol{\zeta}(x)\rangle \varrho(x)\, dx$$

since $\sup_{x\in\mathbf{R}^d} |\varphi(s|x|)\nabla\phi(x)| < \infty$ and $\varphi(s|x|) \longrightarrow 1$ for all $x \in \mathbf{R}^d$ as $s \to 0$.

For the left-hand side of (3.25) we find

$$\begin{aligned} \int_{\mathbf{R}^d} f(x)\Delta\psi_s(x)\, dx = s&\int_{\mathbf{R}^d} f(x)\varphi'(s|x|) \left[2\frac{x}{|x|}\cdot\nabla\phi(x) + \frac{d-1}{|x|}\phi(x) \right] dx \\ &+ s^2\int_{\mathbf{R}^d} f(x)\varphi''(s|x|)\phi(x)\, dx \\ &+ \int_{\mathbf{R}^d} f(x)g(x)\,\varphi(s|x|)\, dx. \end{aligned} \tag{3.27}$$

Recall that $\varphi(s|x|) = 1$ if $|x| \leqslant 1/s$. Since g has compact support, the last integral in (3.27) equals $\int_{\mathbf{R}^d} fg$ for s small enough. If $d \geqslant 3$, then we estimate

$$\begin{aligned} \left| s^2\int_{\mathbf{R}^d} f(x)\varphi''(s|x|)\phi(x)\, dx \right| &\leqslant s^2\|\varphi''\|_{\mathscr{L}^\infty(\mathbf{R})} \|\phi\|_{\mathscr{L}^\infty(\mathbf{R}^d)} \int_{|x|\leqslant 2/s} |f(x)|\, dx \\ &\leqslant Cs^2 s^{-d/q'} \|f\|_{\mathscr{L}^q(\mathbf{R}^d)} \longrightarrow 0. \end{aligned}$$

The constant $C < \infty$ depends on d and q, but not on s. Recall that $q' > d$ and that we can choose q' arbitrarily close to d. For $d = 2$ we obtain a similar estimate with an additional factor $\log(s^{-1})$ because ϕ grows logarithmically. For $d = 1$ we obtain an additional factor s^{-1} because ϕ grows linearly. Similarly, we have

$$\left| s \int_{\mathbf{R}^d} f(x)\varphi'(s|x|) \left[2\frac{x}{|x|} \cdot \nabla\phi(x) + \frac{d-1}{|x|}\phi(x) \right] dx \right| \leqslant C' s s^{-d/q'} \|f\|_{\mathscr{L}^q(\mathbf{R}^d)} \longrightarrow 0.$$

Again the constant $C' < \infty$ depends on d and q, but not on s. Note that $1/|x| \leqslant s$ whenever $\varphi'(s|x|) \neq 0$, and that $\phi(x)/|x|$ is bounded for large x, even for $d = 1$ when $\phi(x)$ grows linearly at infinity. This yields the identity

$$\int_{\mathbf{R}^d} fg = -\int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}, \nabla\phi \rangle \varrho. \tag{3.28}$$

Since $\nabla\phi$ is bounded, we have that $\nabla\phi \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$. We even have $\nabla\phi \in \mathscr{C}_{\varrho,\mathrm{id}}$ because $\Delta\phi = g \geqslant 0$. By choice of $\boldsymbol{\zeta}$, the last term in (3.28) must then be nonnegative. Since $g \in \mathscr{D}(\mathbf{R}^d)$ with $g \geqslant 0$ was arbitrary, we get that $f \geqslant 0$.

Step 2. Formally, the identity (3.14) follows by testing (3.25) against the function $\frac{1}{2}|\mathrm{id}|^2$ and integrating by parts. Since we do not know yet that f is integrable, we must proceed carefully and prove that all integrals are well-defined. To this end, let φ be given by (3.26) and define $\psi_s \in \mathscr{D}(\mathbf{R}^d)$ for all $s > 0$ by

$$\psi_s(x) := |x|^2 \varphi(s|x|) \quad \text{for all } x \in \mathbf{R}^d, \tag{3.29}$$

so that $\nabla\psi_s(x) = \phi_s(x)x$ with

$$\phi_s(x) := 2\varphi(s|x|) + s|x|\varphi'(s|x|) \quad \text{for all } x \in \mathbf{R}^d. \tag{3.30}$$

Note that

$$\begin{aligned} x \cdot \nabla\phi_s(x) &= 3s|x|\varphi'(s|x|) + s^2|x|^2\varphi''(s|x|) \\ &= s\tfrac{\partial}{\partial s}\phi_s(x) \quad \text{for all } x \in \mathbf{R}^d \text{ and } s > 0. \end{aligned}$$

We now test (3.25) against ψ_s and integrate by parts. We obtain

$$dh(s) + s\tfrac{d}{ds}h(s) = r(s) \quad \text{for all } s > 0, \tag{3.31}$$

with

$$h(s) := \int_{\mathbf{R}^d} \phi_s(x) f(x)\,dx \quad \text{and} \quad r(s) := -\int_{\mathbf{R}^d} \phi_s(x) \langle x, \boldsymbol{\zeta}(x) \rangle \varrho(x)\,dx.$$

Note that h and r are smooth for all $s > 0$.

The function ϕ_s is bounded in $\mathscr{L}^\infty(\mathbf{R}^d)$ uniformly in s, and $\phi_s(x) \longrightarrow 2$ for all $x \in \mathbf{R}^d$ as $s \to 0$. By dominated convergence, we therefore obtain

$$r(s) \longrightarrow -2\int_{\mathbf{R}^d} \langle x, \boldsymbol{\zeta}(x) \rangle \varrho(x)\,dx =: r_0 \quad \text{as } s \to 0. \tag{3.32}$$

On the other hand, from (3.31) we find that

$$\alpha^d h(\alpha) = \beta^d h(\beta) + \int_\beta^\alpha s^{d-1} r(s)\,ds \quad \text{for all } 0 < \beta < \alpha.$$

We claim that $\beta^d h(\beta) \longrightarrow 0$ as $\beta \to 0$. Indeed, since $f \in \mathscr{L}^q(\mathbf{R}^d)$, we estimate

$$\beta^d |h(\beta)| \leqslant C\beta^d \int_{|x| \leqslant 2/\beta} |f(x)|\,dx \leqslant C'\beta^d \beta^{-d/q'} \|f\|_{\mathscr{L}^q(\mathbf{R}^d)} \longrightarrow 0,$$

where $C := \sup_{s>0} |2\varphi(s) + s\varphi'(s)| < \infty$ and $C' < \infty$ depends on d, C, and q, but not on β. Using (3.32), we then conclude that the function

$$h(\alpha) = \alpha^{-d} \int_0^\alpha s^{d-1} r(s)\, ds \longrightarrow r_0/d \quad \text{as } \alpha \to 0.$$

To show that f is integrable, we now apply a similar argument again: Let

$$k(s) := \int_{\mathbf{R}^d} \varphi(s|x|) f(x)\, dx \quad \text{for all } s > 0.$$

By the definition of ϕ_s and $h(s)$ above, we have that

$$2k(s) + s\tfrac{d}{ds}k(s) = h(s) \quad \text{for all } s > 0,$$

which implies that

$$\alpha^2 k(\alpha) = \beta^2 k(\beta) + \int_\beta^\alpha s h(s)\, ds \quad \text{for all } 0 < \beta < \alpha.$$

We claim that $\beta^2 k(\beta) \longrightarrow 0$ as $\beta \to 0$. Indeed, since $f \in \mathscr{L}^q(\mathbf{R}^d)$, we estimate

$$\beta^2 k(\beta) \leqslant 2\beta^2 \int_{|x| \leqslant 2/\beta} |f(x)|\, dx \leqslant C'' \beta^2 \beta^{-d/q'} \|f\|_{\mathscr{L}^q(\mathbf{R}^d)} \longrightarrow 0.$$

where the constant $C'' < \infty$ depends on d and q, but not on β. Recall that we may choose q' arbitrarily close to d, thus d/q' close to 1. Hence

$$k(\alpha) = \alpha^{-2} \int_0^\alpha s h(s)\, ds \longrightarrow r_0/2d \quad \text{as } \alpha \to 0.$$

On the other hand, by monotone convergence we know that

$$k(\alpha) \longrightarrow \int_{\mathbf{R}^d} f(x)\, dx \quad \text{as } \alpha \to 0.$$

Then equation (3.14) follows. In particular, we have that $f \in \mathscr{L}^1(\mathbf{R}^d)$.

Step 3. If $d = 1$, then there is nothing left to do, so we assume now that $d \geqslant 2$. Consider a general $B \in \mathcal{S}^d_{++}$ with $\operatorname{tr}(B^2) = d$. Then we claim that

$$\mathbf{r} \in \mathscr{C}_{\varrho,B} \iff \bar{\mathbf{r}} \in \mathscr{C}_{\bar\varrho,\mathrm{id}}, \tag{3.33}$$

$$\boldsymbol{\zeta} \in \mathscr{C}^\perp_{\varrho,B} \iff \bar{\boldsymbol{\zeta}} \in \mathscr{C}^\perp_{\bar\varrho,\mathrm{id}}, \tag{3.34}$$

where

$$\bar{\mathbf{r}} := B\mathbf{r} \circ \ell_B, \quad \bar{\boldsymbol{\zeta}} := B^{-1}\boldsymbol{\zeta} \circ \ell_B, \quad \text{and} \quad \bar\varrho := \det(B)\varrho \circ \ell_B. \tag{3.35}$$

Indeed if $\mathbf{r} \in \mathscr{C}_{\varrho,B}$ and $\bar{\mathbf{r}}, \bar\varrho$ are defined by (3.35), then

$$\int_{\mathbf{R}^d} |\bar{\mathbf{r}}(y)|^2 \bar\varrho(y)\, dy = \int_{\mathbf{R}^d} |B\mathbf{r}(x)|^2 \varrho(x)\, dx,$$

which is equivalent to $\int_{\mathbf{R}^d} |\mathbf{r}|^2 \varrho$ since B is strictly positive definite. By the definition of $\mathscr{C}_{\varrho,B}$, there exists a distribution T such that $[\mathbf{r}] = \nabla T$ in $\mathscr{D}'(\mathbf{R}^d)$. Recall that $\mathbf{r}$

is defined almost everywhere in $\mathbf{R}^d$ since $\varrho(x) > 0$ for a.e. $x \in \mathbf{R}^d$. Then

$$\begin{aligned}
&\int_{\mathbf{R}^d} \langle \bar{\mathbf{r}}(y), \xi(y)\rangle\,dy \\
&\qquad = \int_{\mathbf{R}^d} \left\langle \mathbf{r}(x), \frac{B\xi(B^{-1}x)}{\det(B)} \right\rangle dx \\
&\qquad = \nabla T\bigg(\frac{B\xi \circ \ell_{B^{-1}}}{\det(B)} \bigg) \\
&\qquad = -T\bigg(\frac{\nabla \cdot (B\xi \circ \ell_{B^{-1}})}{\det(B)} \bigg) = -T\bigg(\frac{(\nabla \cdot \xi) \circ \ell_{B^{-1}}}{\det(B)} \bigg) = \Big(\nabla (T \circ \ell_B) \Big)(\xi)
\end{aligned}$$

for all $\xi \in \mathscr{D}(\mathbf{R}^d)$, and thus $[\bar{\mathbf{r}}] = \nabla(T \circ \ell_B)$ in $\mathscr{D}'(\mathbf{R}^d)$. By (3.10), we have

$$\Delta(T \circ \ell_B) = \Big(\langle B^2, D^2 \rangle T \Big) \circ \ell_B \quad \text{in } \mathscr{D}'(\mathbf{R}^d),$$

which is nonnegative if $\langle B^2, D^2\rangle T \geqslant 0$ in $\mathscr{D}'(\mathbf{R}^d)$. Hence $\mathbf{r} \in \mathscr{C}_{\varrho,B}$ entails $\bar{\mathbf{r}} \in \mathscr{C}_{\bar{\varrho},\mathrm{id}}$. Since B is invertible, these arguments work both ways, and so (3.33) follows.

For any $\boldsymbol{\zeta} \in \mathscr{C}_{\varrho,B}^\perp$ and $\bar{\mathbf{r}} \in \mathscr{C}_{\bar{\varrho},\mathrm{id}}$ let $\bar{\boldsymbol{\zeta}}, \mathbf{r}$ be defined by (3.35). Then

$$\int_{\mathbf{R}^d} |\bar{\boldsymbol{\zeta}}(y)|^2 \bar{\varrho}(y)\,dy = \int_{\mathbf{R}^d} |B^{-1}\boldsymbol{\zeta}(x)|^2 \varrho(x)\,dx,$$

which is equivalent to $\int_{\mathbf{R}^d} |\boldsymbol{\zeta}|^2 \varrho$ since B is strictly positive definite. We have

$$\int_{\mathbf{R}^d} \langle \bar{\boldsymbol{\zeta}}(y), \bar{\mathbf{r}}(y)\rangle \bar{\varrho}(y)\,dy = \int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}(x), \mathbf{r}(x)\rangle \varrho(x)\,dx,$$

which is nonnegative since $\mathbf{r} \in \mathscr{C}_{\varrho,B}^\perp$, by (3.33). We conclude that $\bar{\boldsymbol{\zeta}} \in \mathscr{C}_{\bar{\varrho},\mathrm{id}}^\perp$. Again the argument works both ways since B is invertible, which proves (3.34).

For given $\boldsymbol{\zeta} \in \mathscr{C}_{\varrho,B}^\perp$ let $\bar{\boldsymbol{\zeta}} \in \mathscr{C}_{\bar{\varrho},\mathrm{id}}^\perp$ be defined by (3.35). As shown in Step 2, there exists a nonnegative $\bar{f} \in \mathscr{L}^1(\mathbf{R}^d)$ with the property that

$$\begin{gathered}
\Delta \bar{f} = \nabla \cdot (\bar{\varrho}\bar{\boldsymbol{\zeta}}) \quad \text{in } \mathscr{D}'(\mathbf{R}^d), \\
d \int_{\mathbf{R}^d} \bar{f}(y)\,dy = -\int_{\mathbf{R}^d} \langle y, \bar{\boldsymbol{\zeta}}(y)\rangle \bar{\varrho}(y)\,dy.
\end{gathered}$$

Recall that the $\mathscr{L}^2(\mathbf{R}^d, \bar{\varrho})$-norm of $\bar{\boldsymbol{\zeta}}$ can be bounded by the $\mathscr{L}^2(\mathbf{R}^d, \varrho)$-norm of $\boldsymbol{\zeta}$ since B is invertible. We now define $f \in \mathscr{L}^1(\mathbf{R}^d)$ by $f := \bar{f} \circ \ell_{B^{-1}} / \det(B)$, which is again a nonnegative function. Then we obtain the identity

$$\begin{aligned}
&\int_{\mathbf{R}^d} f(x) \langle B^2, D^2\varphi(x)\rangle\,dx \\
&\qquad = \int_{\mathbf{R}^d} \bar{f}(y) \Big(\langle B^2, D^2\varphi\rangle(By) \Big)\,dy = \int_{\mathbf{R}^d} \bar{f}(y) \Big(\Delta(\varphi \circ \ell_B) \Big)(y)\,dy
\end{aligned}$$

for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$. Note that $\varphi \circ \ell_B \in \mathscr{D}(\mathbf{R}^d)$. Then (2.14) implies that

$$\begin{aligned}
&\int_{\mathbf{R}^d} \bar{f}(y) \Big(\Delta(\varphi \circ \ell_B) \Big)(y)\,dy \\
&\qquad = -\int_{\mathbf{R}^d} \Big\langle \bar{\boldsymbol{\zeta}}(y), \Big(\nabla(\varphi \circ \ell_B) \Big)(y) \Big\rangle \bar{\varrho}(y)\,dy \\
&\qquad = -\int_{\mathbf{R}^d} \Big\langle \bar{\boldsymbol{\zeta}}(y), B(\nabla\varphi)(By) \Big\rangle \bar{\varrho}(y)\,dy = -\int_{\mathbf{R}^d} \langle \boldsymbol{\zeta}(x), \nabla\varphi(x)\rangle \varrho(x)\,dx
\end{aligned}$$

for all $\varphi \in \mathscr{D}(\mathbf{R}^d)$, from which identity (3.13) follows.

To prove (3.14), we test (3.13) against the function ψ_s defined in (3.29). Then

$$\int_{\mathbf{R}^d} f(x)\langle B^2, D^2\psi_s(x)\rangle\,dx = d\int_{\mathbf{R}^d} f(x)\phi_s(x)\,dx + \int_{\mathbf{R}^d} f(x)\bar{\phi}_s(x)\frac{|Bx|^2}{|x|^2}\,dx,$$

where ϕ_s is defined in (3.30) and

$$\bar{\phi}_s(x) := 3s|x|\varphi'(s|x|) + s^2|x|^2\varphi''(s|x|) \quad \text{for all } x \in \mathbf{R}^d.$$

Both ϕ_s and $\bar{\phi}_s$ are uniformly bounded in $x \in \mathbf{R}^d$ and $s > 0$. Since $f \in \mathscr{L}^1(\mathbf{R}^d)$ and since $\phi_s(x) \longrightarrow 2$ and $\bar{\phi}_s(x) \longrightarrow 0$ for all $x \in \mathbf{R}^d$ as $s \to 0$, we can then apply the dominated convergence theorem to conclude that

$$\int_{\mathbf{R}^d} f(x)\langle B^2, D^2\psi_s(x)\rangle\,dx \longrightarrow 2d\int_{\mathbf{R}^d} f(x)\,dx \quad \text{as } s \to 0.$$

For the right-hand side of (3.13)/(3.14), we proceed as in Step 2 above. □

4. Variational Time Discretization

In this section, we discuss in detail a variational time discretization for the system of isentropic Euler equations (1.1). This scheme is a variation of the one we proposed in [12]. In the next two sections, we will then show that a sequence of approximate solutions generated by this discretization converges weakly and generates a Young measure that is a measure-valued solution of the isentropic Euler equations.

Definition 4.1 (Energy)**.** Let $U\colon [0,\infty) \longrightarrow \mathbf{R}$ be a proper, lower semicontinuous, convex function with $U(0) = 0$ such that the map $r \mapsto r^d U(r^{-d})$ is strictly convex and nonincreasing on $(0,\infty)$. We also assume for simplicity that U is nonnegative. For $\varrho \in \mathscr{P}(\mathbf{R}^d)$ we define the *internal energy* as

$$\mathcal{U}(\varrho) := \begin{cases} \displaystyle\int_{\mathbf{R}^d} U(\varrho(x))\,dx & \text{if } \varrho \ll \mathcal{L}^d, \\ +\infty & \text{otherwise.} \end{cases}$$

For $\varrho \in \mathscr{P}(\mathbf{R}^d)$ and vector-valued $\mathbf{m} \in \mathscr{M}_{\mathrm{loc}}(\mathbf{R}^d)$ we define the *kinetic energy* as

$$\mathcal{K}(\varrho, \mathbf{m}) := \begin{cases} \displaystyle\int_{\mathbf{R}^d} \tfrac{1}{2}|\mathbf{u}(x)|^2\,\varrho(dx) & \text{if } \mathbf{m} = \varrho\mathbf{u} \text{ with } \mathbf{u} \in \mathscr{L}^2(\mathbf{R}^d, \varrho), \\ +\infty & \text{otherwise.} \end{cases}$$

We define the *total energy* as the sum $\mathcal{E}(\varrho, \mathbf{m}) := \mathcal{K}(\varrho, \mathbf{m}) + \mathcal{U}(\varrho)$. Occasionally, we will write $\mathcal{E}(\varrho, \mathbf{u})$ instead of $\mathcal{E}(\varrho, \mathbf{m})$ with $\mathbf{m} = \varrho\mathbf{u}$, to simplify notation.

Definition 4.2 (Minimal Acceleration Cost)**.** For any $s > 0$ let

$$A_s(\mathbf{x}_1, \mathbf{x}_2)^2 := 3\left|\frac{x_2 - x_1}{\tau} - \frac{\xi_2 + \xi_1}{2}\right|^2 + \frac{1}{4}|\xi_2 - \xi_1|^2 \tag{4.1}$$

for all $\mathbf{x}_i \equiv (x_i, \xi_i) \in \mathbf{R}^{2d}$. Then the Minimal Acceleration Cost is the functional

$$\mathbf{A}_\tau(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)^2 := \inf\left\{ \iint_{\mathbf{R}^{2d}\times\mathbf{R}^{2d}} A_\tau(\mathbf{x}_1, \mathbf{x}_2)^2\,\boldsymbol{\beta}(d\mathbf{x}_1, d\mathbf{x}_2) \colon \boldsymbol{\beta} \in \Gamma(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) \right\}, \tag{4.2}$$

defined for all measures $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathscr{P}(\mathbf{R}^{2d})$. Here $\Gamma(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)$ is the set of all transport plans $\boldsymbol{\beta} \in \mathscr{P}(\mathbf{R}^{2d}\times\mathbf{R}^{2d})$ such that $\pi^i_{\#}\boldsymbol{\beta} = \boldsymbol{\mu}^i$, where $\pi^i\colon \mathbf{R}^{2d}\times\mathbf{R}^{2d} \longrightarrow \mathbf{R}^{2d}$ denotes the projection onto the ith factor with $i = 1\ldots 2$.

For given $\tau > 0$, consider now $\varrho \in \mathscr{P}(\mathbf{R}^d)$ and $\mathbf{u} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ with the property that $\mathcal{U}(\varrho) < \infty$ and $\hat{\mathbf{r}}_\tau := \mathrm{id} + \tau\mathbf{u} \in \mathscr{C}_\varrho$. Define a measure $\boldsymbol{\mu} \in \mathscr{P}(\mathbf{R}^{2d})$ by

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \boldsymbol{\mu}(d\mathbf{x}) := \int_{\mathbf{R}^d} \varphi(x, \mathbf{u}(x))\varrho(x)\, dx \quad \text{for all } \varphi \in \mathscr{C}_\mathrm{b}(\mathbf{R}^{2d}), \tag{4.3}$$

with $\mathbf{x} \equiv (x, \xi)$. We proved in [12] that for all $s \in (0, \tau)$ there exists a unique

$$\boldsymbol{\mu}_s \in \operatorname{argmin} \Big\{ \mathbf{A}_s(\boldsymbol{\mu}, \boldsymbol{\mu}^*)^2 + \mathcal{U}(\pi_{\#}\boldsymbol{\mu}^*) \colon \boldsymbol{\mu}^* \in \mathscr{P}(\mathbf{R}^{2d}) \Big\}, \tag{4.4}$$

where $\pi \colon \mathbf{R}^d \times \mathbf{R}^d \longrightarrow \mathbf{R}^d$ denotes the projection onto the first factor. The minimizer $\boldsymbol{\mu}_s$ is induced by a density $\varrho_s \in \mathscr{P}(\mathbf{R}^d)$ and velocity $\mathbf{u}_s \in \mathscr{L}^2(\mathbf{R}^d, \varrho_s)$ as in equation (4.3). In fact, the new density ϱ_s is the unique

$$\varrho_s \in \operatorname{argmin} \left\{ \frac{3}{4s^2} \mathbf{W}(\hat{\varrho}_s, \varrho^*)^2 + \mathcal{U}(\varrho^*) \colon \varrho^* \in \mathscr{P}(\mathbf{R}^d) \right\}, \tag{4.5}$$

where $\hat{\varrho}_s := (\hat{\mathbf{r}}_s)_{\#}\varrho$ with $\hat{\mathbf{r}}_s := \mathrm{id} + s\mathbf{u}$ is the push-forward measure. Notice that since ϱ is absolutely continuous with respect to the Lebesgue measure, so is $\hat{\varrho}_s$; see Proposition 1.3 in [19]. Moreover, there is a unique optimal transport map pushing $\hat{\varrho}_s$ forward to ϱ, which is given by $\hat{\mathbf{z}}_s := \hat{\mathbf{r}}_s^{-1}$; see Lemma 7.2.1 in [5]. Since ϱ_s in (4.5) depends continuously on $s > 0$ and $\hat{\varrho}_s$ (see Lemma 3.2.1 in [5]), and

$$\mathbf{W}(\hat{\varrho}_{s_1}, \hat{\varrho}_{s_2}) = |s_1 - s_2| \|\mathbf{u}\|_{\mathscr{L}^2(\mathbf{R}^d, \varrho)} \quad \text{for all } s_1, s_2 \in [0, \tau]$$

because of Theorem 7.2.2 in [5], the measure ϱ_s depends continuously on $s > 0$. In particular, we have that $\varrho_\tau = \lim_{s \to \tau} \varrho_s$ with respect to the Wasserstein distance. There is a unique optimal transport map $\mathbf{r}_s$ pushing ϱ_s forward to $\hat{\varrho}_s$ given by

$$\mathbf{r}_s := \mathrm{id} + \frac{2s^2}{3} \nabla U'(\varrho_s) \quad \text{for all } s \in (0, \tau), \tag{4.6}$$

which is invertible since $\hat{\varrho}_s$ is absolutely continuous with respect to the Lebesgue measure. We denote the inverse map by $\mathbf{z}_s := \mathbf{r}_s^{-1}$. Equation (4.6) is still correct if $s = \tau$, but then $\mathbf{r}_\tau$ may no longer be invertible anymore. We have

$$\varrho_s \mathcal{L}^d = \left(\Big(\mathrm{id} + s\mathbf{u} \Big)^{-1} \circ \left(\mathrm{id} + \frac{2s^2}{3} \nabla U'(\varrho_s) \right) \right)^{-1}_{\#} (\varrho \mathcal{L}^d), \tag{4.7}$$

$$\mathbf{u}_s = \mathbf{u} \circ \Big(\mathrm{id} + s\mathbf{u} \Big)^{-1} \circ \left(\mathrm{id} + \frac{2s^2}{3} \nabla U'(\varrho_s) \right) - \tau \nabla U'(\varrho_s). \tag{4.8}$$

Moreover, we have the following energy dissipation estimate:

$$\mathcal{E}(\varrho_s, \mathbf{u}_s) + \frac{s^2}{6} \int_{\mathbf{R}^d} |\nabla U'(\varrho_s)|^2 \varrho_s \, dx \leqslant \mathcal{E}(\varrho, \mathbf{u}). \tag{4.9}$$

We refer the reader to [12] for more details.

For all $s \in (0, \tau)$ we now define a plan $\boldsymbol{\alpha}_s \in \mathscr{P}(\mathbf{R}^{3d})$ by

$$\int_{\mathbf{R}^{3d}} \varphi(\mathbf{x}, y)\, \boldsymbol{\alpha}_s(d\mathbf{x}, dy) := \int_{\mathbf{R}^d} \varphi\Big(x, \mathbf{u}(x), \mathbf{z}_s\big(\hat{\mathbf{r}}_s(x)\big)\Big)\, \varrho(dx)$$

for all $\varphi \in \mathscr{C}_\mathrm{b}(\mathbf{R}^{3d})$, with $\mathbf{x} \equiv (x, \xi)$. Then we have the following identities:

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \boldsymbol{\mu}(d\mathbf{x}) = \int_{\mathbf{R}^{3d}} \varphi(\mathbf{x})\, \boldsymbol{\alpha}_s(d\mathbf{x}, dy), \tag{4.10}$$

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \boldsymbol{\mu}_s(d\mathbf{x}) = \int_{\mathbf{R}^{3d}} \varphi(y, \beta(\mathbf{x}, y))\, \boldsymbol{\alpha}_s(d\mathbf{x}, dy)$$

for all $\varphi \in \mathscr{C}_\mathrm{b}(\mathbf{R}^{2d})$, and

$$\int_{\mathbf{R}^d} |\nabla U'(\varrho_s(x))|^2 \varrho_s(x)\, dx = \int_{\mathbf{R}^{3d}} |\beta_s(\mathbf{x},y) - \xi|^2\, \boldsymbol{\alpha}_s(d\mathbf{x},dy),$$

where $\beta_s(\mathbf{x},y) := \xi - \frac{3}{2s}((x+s\xi)-y)$ for all $\mathbf{x}\in\mathbf{R}^{2d}$ and $y\in\mathbf{R}^d$. Again we refer the reader to [12] (in particular Proposition 4.5) for further information.

We now claim that the sequence $\{\boldsymbol{\alpha}_s\}$ is tight: Note that the first two marginals of $\boldsymbol{\alpha}_s$ do not depend on s; see (4.10). The third maginal of $\boldsymbol{\alpha}_s$ equals ϱ_s, and $\{\varrho_s\}$ converges with respect to the Wasserstein distance to ϱ_τ as $s\to\tau$. Therefore $\{\varrho_s\}$ is a tight set. The tightness of $\{\boldsymbol{\alpha}_s\}$ then follows from Lemma 5.2.2 in [5]. Applying Prokhorov's theorem again, we conclude that there exists a subsequence (which we still denote by $\{\boldsymbol{\alpha}_s\}$ for simplicity) and a measure $\boldsymbol{\alpha}_\tau\in\mathscr{P}(\mathbf{R}^{3d})$ such that

$$\boldsymbol{\alpha}_s \rightharpoonup \boldsymbol{\alpha}_\tau \quad \text{narrowly as } s\to\tau.$$

By Proposition 7.1.3 in [5], the measure $\hat{\boldsymbol{\gamma}}_\tau\in\mathscr{P}(\mathbf{R}^{2d})$ defined by

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \hat{\boldsymbol{\gamma}}_\tau(d\mathbf{x}) := \int_{\mathbf{R}^{3d}} \varphi(x, x+\tau\xi)\, \boldsymbol{\alpha}_\tau(d\mathbf{x},dy) \quad \text{for all } \varphi\in\mathscr{C}_\mathrm{b}(\mathbf{R}^{2d})$$

is an optimal transport plan between ϱ and $\hat{\varrho}_\tau$, which is induced by the map $\hat{\mathbf{r}}_\tau$ and thus uniquely determined. Similarly, the measure $\boldsymbol{\gamma}_\tau\in\mathscr{P}(\mathbf{R}^{2d})$ defined by

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \boldsymbol{\gamma}_\tau(d\mathbf{x}) := \int_{\mathbf{R}^{3d}} \varphi(y, x+\tau\xi)\, \boldsymbol{\alpha}_\tau(d\mathbf{x},dy) \quad \text{for all } \varphi\in\mathscr{C}_\mathrm{b}(\mathbf{R}^{2d})$$

is an optimal transport plan between ϱ_τ and $\hat{\varrho}_\tau$, induced by $\mathbf{r}_\tau$ given by (4.6).

We now define a measure $\boldsymbol{\mu}_\tau\in\mathscr{P}(\mathbf{R}^{2d})$ by

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \boldsymbol{\mu}_\tau(d\mathbf{x}) := \int_{\mathbf{R}^{3d}} \varphi(y, \beta_\tau(\mathbf{x},y))\, \boldsymbol{\alpha}_\tau(d\mathbf{x},dy) \quad \text{for all } \varphi\in\mathscr{C}_\mathrm{b}(\mathbf{R}^{2d}). \tag{4.11}$$

By lower semicontinuity, we then obtain from (4.9) the energy inequality

$$\bigg(\int_{\mathbf{R}^{2d}} \tfrac12|\xi|^2\, \boldsymbol{\mu}_\tau(d\mathbf{x}) + \mathcal{U}(\varrho_\tau) \bigg) + \frac{\tau^2}{6}\int_{\mathbf{R}^d} |\nabla U'(\varrho_\tau)|^2 \varrho_\tau\, dx \leqslant \mathcal{E}(\varrho,\mathbf{u}). \tag{4.12}$$

Note that the new state $\boldsymbol{\mu}_\tau$ may no longer be monokinetic. That is, the measure $\boldsymbol{\mu}_\tau$ may not be induced by the density ϱ_τ and an Eulerian velocity field as in (4.3). If we now apply the velocity projection defined in Theorem 2.10 to $\boldsymbol{\mu}_\tau$, however, then we obtain a velocity $\mathbf{u}_\tau\in\mathscr{L}^2(\mathbf{R}^d,\varrho_\tau)$ with the property that $\mathrm{id}+\tau\mathbf{u}_\tau\in\mathscr{C}_{\varrho_\tau}$. Moreover, the kinetic energy of $(\varrho_\tau,\mathbf{u}_\tau)$ is bounded above by the old energy:

$$\int_{\mathbf{R}^d} \tfrac12|\mathbf{u}_\tau(x)|^2\varrho_\tau(x)\, dx \leqslant \int_{\mathbf{R}^{2d}} \tfrac12|\xi|^2 \boldsymbol{\mu}_\tau(d\mathbf{x}). \tag{4.13}$$

In fact, we have a more precise estimate, which gives a lower bound bound on the energy dissipation; see (2.15). Combining (4.13) with (4.12), we obtain

$$\mathcal{E}(\varrho_\tau,\mathbf{u}_\tau) + \frac{\tau^2}{6}\int_{\mathbf{R}^d} |\nabla U'(\varrho_\tau)|^2\varrho_\tau\, dx \leqslant \mathcal{E}(\varrho,\mathbf{u}). \tag{4.14}$$

Remark 4.3. In [12] we chose the timestep $\tau>0$ in such a way that the pushforward measure $\hat{\varrho}_\tau$ is absolutely continuous with respect to the Lebesgue measure, and so the plan $\boldsymbol{\alpha}_\tau$ between ϱ, $\hat{\varrho}_\tau$, and ϱ_τ is uniquely determined by the two optimal and *invertible* transport maps $\hat{\mathbf{r}}_\tau$ and $\mathbf{r}_\tau$. Here we use the velocity projection to compute an optimal transport velocity corresponding to some *fixed* timestep $\tau>0$ and then push the density forward in the direction of the new velocity as far as possible. This typically creates singular measures. In fact, in a very interesting, recent paper [21], the authors explore the connection between the velocity projection and the sticky particle dynamics for one-dimensional pressureless flows, which are modeled by the system of conservation laws (1.1) with $P(r) = 0$ for all $r \geqslant 0$. Then the density may become a singular measure in finite time. The physical interpretation is the following: Due to the absence of pressure, fluid element do not notice each other unless they collide, in which case they stick together to form a bigger particle. Total mass and momentum is preserved during the collision, but some kinetic energy may get lost. The velocity projection captures this behavior since in regions where the given transport map is not optimal, it introduces constant segments. This implies that a positive amount of mass may be concentrated at some location and thereby form a Dirac measure. We refer the reader to [21] for further details.

Note that even if the optimal transport plans between ϱ and $\hat{\varrho}_\tau$ and between ϱ_τ and $\hat{\varrho}_\tau$ are unique and induced by the maps $\hat{\mathbf{r}}_\tau$ resp. $\mathbf{r}_\tau$, when $\hat{\varrho}_\tau$ is singular there may be many plans $\boldsymbol{\alpha}_\tau$ connecting the three measures. But the precise trajectory of fluid elements determines the new velocity through the function β_τ above, which depends on the old position/velocity and the new position. All these plans give the same value for the minimal acceleration cost since $\mathbf{A}_\tau$ reduces to the Wasserstein distance in (4.5) from the pushforward measure $\hat{\varrho}_\tau$; see Proposition 4.5 in [12]. We therefore choose for $\boldsymbol{\alpha}_\tau$ the limit of $\boldsymbol{\alpha}_s$ as $s \to \tau$. Notice that when the internal energy vanishes, then the minimization (4.4) reduces to a free transport. That is, we have $\boldsymbol{\mu}_\tau = (F_\tau)_\# \boldsymbol{\mu}$, where $F_\tau(\mathbf{x}) := (x + \tau\xi, \xi)$ for all $\mathbf{x} \equiv (x, \xi) \in \mathbf{R}^{2d}$.

Definition 4.4 (Time Discretization)**.** Let $\tau > 0$ and $(\bar{\varrho}, \bar{\mathbf{u}}) \in \mathbb{T}\mathscr{P}(\mathbf{R}^d)$ be given. Then we define approximate solutions $(\varrho_k^\tau, \mathbf{u}_k^\tau) \in \mathbb{T}\mathscr{P}(\mathbf{R}^d)$ at discrete times $t_k^\tau := k\tau$ for all $k \in \mathbf{N} \cup \{0\}$ by executing the steps of the following program:

(1) **Initial Data**
Let $k = 0$ and $\varrho_0^\tau := \bar{\varrho}$. Define a measure $\boldsymbol{\mu}_0^\tau \in \mathscr{P}(\mathbf{R}^{2d})$ by

$$\int_{\mathbf{R}^{2d}} \varphi(\mathbf{x})\, \boldsymbol{\mu}_0^\tau(d\mathbf{x}) := \int_{\mathbf{R}^d} \varphi(x, \bar{\mathbf{u}}(x)) \bar{\varrho}(x)\, dx \quad \text{for all } \varphi \in \mathscr{C}_\mathrm{b}(\mathbf{R}^{2d}).$$

(2) **Velocity Projection**
Let $\mathbf{u}_k^\tau \in \mathscr{L}^2(\mathbf{R}^d, \varrho_k^\tau)$ be the velocity obtained by the velocity projection defined in Theorem 2.10 corresponding to the measure $\boldsymbol{\mu} = \boldsymbol{\mu}_k^\tau$.

(3) **Energy Minimization**
Let $\boldsymbol{\alpha}_s \in \mathscr{P}(\mathbf{R}^{3d})$ be the plan defined above for $\boldsymbol{\mu} = \boldsymbol{\mu}_k^\tau$ and $s \in (0, \tau)$. We denote by $\boldsymbol{\alpha}_{k+1}^\tau$ some narrow limit of $\boldsymbol{\alpha}_s$ as $s \to \tau$, and by $\boldsymbol{\mu}_{k+1}^\tau \in \mathscr{P}(\mathbf{R}^{2d})$ the measure given by (4.11). Increase k by one and continue with (2).

Then we define a piecewise constant curve $(\varrho^\tau, \mathbf{u}^\tau)\colon [0, \infty) \longrightarrow \mathbb{T}\mathscr{P}(\mathbf{R}^d)$ by

$$(\varrho^\tau, \mathbf{u}^\tau)(t) := (\varrho_k^\tau, \mathbf{u}_k^\tau) \quad \text{for all } t \in [t_k^\tau, t_{k+1}^\tau) \text{ and } k \in \mathbf{N} \cup \{0\}.$$

Finally, we define the momentum $\mathbf{m}^\tau(t) := (\varrho^\tau \mathbf{u}^\tau)(t)$ for all $t \in [0, \infty)$.

Combining (2.15) with (4.14), we obtain the following result:

Proposition 4.5 (Energy Inequality)**.** *Let the density/velocity $(\varrho_k^\tau, \mathbf{u}_k^\tau)$ be given by Definition 4.4 for $k \in \mathbf{N} \cup \{0\}$. Then we have the estimate*

$$\begin{aligned}\mathcal{E}(\varrho_{k+1}^\tau, \mathbf{u}_{k+1}^\tau) &+ \frac{\tau^2}{6} \int_{\mathbf{R}^d} |\nabla U'(\varrho_{k+1}^\tau)|^2 \varrho_{k+1}^\tau \, dx \\ &+ \frac{1}{2} \int_{\mathbf{R}^{2d}} |\xi - \mathbf{u}_k^\tau(x)|^2 \, \boldsymbol{\mu}_k^\tau(dx, d\xi) \leqslant \mathcal{E}(\varrho_k^\tau, \mathbf{u}_k^\tau).\end{aligned}$$

5. Young Measures

We consider now a sequence $\tau^n \to 0$ and construct a corresponding sequence of densities/momentums as in Definition 4.4. We will use the superscript n instead of τ^n in the following, to simplify notation. We obtain a sequence of functions

$$(\varrho^n, \mathbf{m}^n) \colon [0, \infty) \times \mathbf{R}^d \longrightarrow \mathbb{H} \quad \text{for all } n \in \mathbf{N},$$

taking values in the set $\mathbb{H} := ((0, \infty) \times \mathbf{R}^d) \cup \{(0, 0)\}$.

Our goal is to show that there exists a subsequence of $\{(\varrho^n, \mathbf{m}^n)\}$ that converges to a solution of the system of isentropic Euler equations (1.1), in a sense to be specified below. Let us first establish precompactness of the sequence in a suitable topology. Consistency with (1.1) will be considered in Section 6.

The only uniform bound on $\{(\varrho^n, \mathbf{m}^n)\}$ that is readily available, is the bound on the total energy provided by Proposition 4.5. We have that

$$\sup_n \operatorname*{ess\,sup}_{t \in [0,\infty)} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t, x) \, dx < \infty, \tag{5.1}$$

with $E(r, m) := \frac{1}{2}|m|^2/r + U(r)$ for all $(r, m) \in \mathbb{H}$. The energy dissipation estimate

$$\begin{aligned}&\sum_{k=1}^\infty \frac{1}{6} (\tau^n)^2 \int_{\mathbf{R}^d} |\nabla U'(\varrho_k^n)|^2 \varrho_k^n \, dx \\ &\qquad \leqslant \sum_{k=1}^\infty \bigg(\int_{\mathbf{R}^d} E(\varrho_{k-1}^n, \mathbf{m}_{k-1}^n) \, dx - \int_{\mathbf{R}^d} E(\varrho_k^n, \mathbf{m}_k^n) \, dx \bigg),\end{aligned}$$

which also follows from Proposition 4.5, is too weak to enforce strong convergence of $\{\varrho^n\}$ in some Lebesgue space. We therefore try to identify a notion of convergence that relies only on the energy bound (5.1). Let us assume for the moment that the internal energy U is given by the power-law (1.3). Then (5.1) implies that

$$\sup_n \operatorname*{ess\,sup}_{t \in [0,\infty)} \int_{\mathbf{R}^d} \Big((\varrho^n)^\gamma + |\mathbf{m}^n|^p \Big)(t, x) \, dx < \infty \tag{5.2}$$

for $\gamma > 1$ and $p := 2\gamma/(\gamma + 1)$. By Banach-Alaoglu theorem, there exists a subsequence (which we still denote by $\{(\varrho^n, \mathbf{m}^n)\}$ for simplicity) such that

$$\begin{aligned}\varrho^n &\rightharpoonup \varrho \quad \text{weak* in } \mathscr{L}^\infty\big([0, \infty), \mathscr{L}^\gamma(\mathbf{R}^d)\big), \\ \mathbf{m}^n &\rightharpoonup \mathbf{m} \quad \text{weak* in } \mathscr{L}^\infty\big([0, \infty), \mathscr{L}^p(\mathbf{R}^d)\big),\end{aligned} \tag{5.3}$$

for suitable limit density/momentum $(\varrho, \mathbf{m})$. By lower semicontinuity of the total energy (see Section 2.6 in [2]), we find that $\mathbf{m}$ is absolutely continuous with respect to the measure $\varrho\mathcal{L}^d$, so that there exists a unique velocity field $\mathbf{u} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ with $\mathbf{m} = \varrho\mathbf{u}$. Moreover, for a.e. $0 \leqslant a < b < \infty$ we have the estimate

$$\int_{[a,b] \times \mathbf{R}^d} E(\varrho, \mathbf{m})(t, x) \, dx \, dt \leqslant \liminf_{n \to \infty} \int_{[a,b] \times \mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t, x) \, dx \, dt. \tag{5.4}$$

For more general internal energies, the above argument can be modified suitably.

Note that by Proposition 4.5, the map

$$t \mapsto \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt \quad \text{for a.e. } t \in [0,\infty) \tag{5.5}$$

is nonincreasing in time, thus a function of bounded variation. Using Helly's theorem and extracting another subsequence if necessary, we may therefore assume that the sequence of maps (5.5) converges pointwise to a bounded, nonincreasing limit function. Using Fatou's lemma in (5.4), we obtain that

$$\int_{[a,b]\times\mathbf{R}^d} E(\varrho, \mathbf{m})(t,x)\,dx\,dt \leqslant \int_{[a,b]} \left(\lim_{n\to\infty} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(x,t)\,dx \right) dt$$

for a.e. $0 \leqslant a < b < \infty$, which implies the estimate

$$\int_{\mathbf{R}^d} E(\varrho, \mathbf{m})(t,x)\,dx \leqslant \lim_{n\to\infty} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx$$

for a.e. $t \in [0,\infty)$. In general, this inequality can be strict. Estimate (5.1) does not rule out the possibility of energy being lost due to the following mechanisms:

(1) Leakage to Infinity. There exists a sequence of subsets of $\mathbf{R}^d$ that carry a certain fraction of the total energy, and that move to infinity as $n \to \infty$.

(2) Concentrations. It is possible that the sequence of total energy densities concentrates energy on a sequence of subsets in $\mathbf{R}^d$ whose Lebesgue measure converges to zero. In particular, it can happen that the sequence $\{E(\varrho^n, \mathbf{m}^n)\}$ converges weak* to a singular measure. On the other hand, the density/momentum ϱ and $\mathbf{m}$ in (5.3) are Lebesgue measurable, so $E(\varrho, \mathbf{m})$ does not have singular parts.

Concentration of energy also occurs when the density ϱ^n converges to zero in some set, while the velocity $\mathbf{u}^n$ grows without bound in such a way that the kinetic energy stays finite. In that case, the weak* limit of $\{E(\varrho^n, \mathbf{m}^n)\}$ in the measure sense might not be absolutely continuous with respect to $\varrho\mathcal{L}^d$.

(3) Oscillations. The sequence $\{(\varrho^n, \mathbf{m}^n)\}$ oscillates wildly as $n \to \infty$.

In order to show that these effects do not occur for the sequence of approximate solutions $\{(\varrho^n, \mathbf{m}^n)\}$ we need stronger information in addition to the bound (5.1). In the following, we will discuss sufficient conditions that prevent concentration effects. We will also introduce a Young measure that is capable of capturing oscillations. It will allow us to represent limits of nonlinear functions of density and momentum, such as those relevant for the momentum equation in (1.1). Since we are interested in distributional solutions of the isentropic Euler equations, for which test functions are compactly supported, we are not concerned about leakage to infinity.

Our goal is to find a representation of the weak limits of sequences of the form $\{F(\varrho^n, \mathbf{m}^n)\}$ for a sufficiently large class of functions F. This class should contain all nonlinear maps relevant to the momentum equation in (1.1). In the following, we will denote by $\mathscr{W}(\mathbb{H})$ the space of all functions defined as

$$F(r,m) := \varphi(r,m)\Big(1 + E(r,m)\Big) + \phi\Big(\frac{m}{|m|}\Big)\frac{|m|^2}{2r} + cU(r) \tag{5.6}$$

for all $(r,m) \in \mathbb{H}$, with weights

$$\varphi \in \mathscr{C}_0(\mathbb{H}), \quad \phi \in \mathscr{C}(S^{d-1}), \quad \text{and} \quad c \in \mathbf{R}.$$

Here $\mathscr{C}_0(\mathbb{H})$ denotes the closure of the space of continuous functions with compact support in $\mathbb{H}$, with respect to the sup-norm. We have the following relations:

$F(r,m)$	$\varphi(r,m)$	$\phi\Big(\frac{m}{\lvert m\rvert}\Big)$	c
ϱ	$r\Big(1+E(r,m)\Big)^{-1}$	0	0
m	$m\Big(1+E(r,m)\Big)^{-1}$	0	0

Note that $E(r,m) \longrightarrow \infty$ superlinearly in r and quadratically in m, as either $r \to 0$ with $m \neq 0$, or as $r + |m| \to \infty$. This implies that in both cases $\varphi \in \mathscr{C}_0(\mathbb{H})$.

We can represent more complicated nonlinear terms as follows:

$F(r,m)$	$\varphi(r,m)$	$\phi\Big(\frac{m}{\lvert m\rvert}\Big)$	c
$\frac{\lvert m\rvert^2}{2r}$	0	1	0
$\frac{m\otimes m}{r}$	0	$\frac{2m\otimes m}{\lvert m\rvert^2}$	0
$U(r)$	0	0	1
$P(r)$	0	0	$\gamma-1$

Here we assumed the equation of state for a polytropic gas; see (1.2). For more general pressures a combination of nonzero c and φ is necessary. In the isothermal case with $U(r) = \kappa r \log r$, for which $P(r) = \kappa r$, we can even choose $c = 0$.

To simplify notation, we will often use the short-hand $\mathbf{z} = (r,m) \in \mathbb{H}$ in the following. We denote by $\mathscr{P}_E(\mathbb{H})$ the space of all probability measures μ on $\mathbb{H}$ with the property that $\int_{\mathbb{H}} E(\mathbf{z})\,\mu(d\mathbf{z}) < \infty$. Let $Q \subset [0,\infty) \times \mathbf{R}^d$ be a compact subset. For any n we then define a weak*-measurable map $\nu^n \in \mathscr{L}^\infty_{\mathrm{w}}(Q, \mathscr{P}_E(\mathbb{H}))$ by

$$\int_Q \int_{\mathbb{H}} \zeta(t,x) F(\mathbf{z})\,\nu^n_{(t,x)}(d\mathbf{z})\,dx\,dt := \int_Q \zeta(t,x) F(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt$$

for all $\zeta \in \mathscr{L}^1(Q)$ and $F \in \mathscr{C}_{\mathrm{b}}(\mathbb{H})$. Note that the sublevels $\{\mathbf{z} \in \mathbb{H}\colon E(\mathbf{z}) \leqslant R\}$ for all $R \geqslant 0$ are compact and convex. Since (5.1) implies that

$$\sup_n \int_Q \int_{\mathbb{H}} E(\mathbf{z})\,\nu^n_{(t,x)}(d\mathbf{z})\,dx\,dt < \infty,$$

we can use Remark 5.1.5 in [5] to conclude that the sequence of measures $\{\nu^n\}$ is tight, thus precompact in the narrow topology, by Prokhorov's theorem. Therefore there exists a subsequence of $\{\nu^n\}$ (which we still denote by $\{\nu^n\}$ for simplicity) and a limit function $\nu \in \mathscr{L}^\infty_{\mathrm{w}}(Q, \mathscr{P}_E(\mathbb{H}))$ such that

$$\lim_{n\to\infty} \int_Q \int_{\mathbb{H}} \zeta(t,x) F(\mathbf{z})\,\nu^n_{(t,x)}(d\mathbf{z})\,dx\,dt = \int_Q \int_{\mathbb{H}} \zeta(t,x) F(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt \tag{5.7}$$

for all $\zeta \in \mathscr{L}^1(Q)$ and $F \in \mathscr{C}_{\mathrm{b}}(\mathbb{H})$. Our goal is to extend this convergence to all F from the class $\mathscr{W}(\mathbb{H})$. To achieve this goal, we need additional assumptions. Our first assumption excludes the possibility of energy concentrating on small sets:

Assumption 5.1 (Uniform Integrability)**.** For all $0 \leqslant a < b < \infty$ we have

$$\lim_{R\to\infty} \int_{[a,b]\times\mathbf{R}^d} \mathbf{1}_{\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n,\mathbf{m}^n)(t,x)\,dx\,dt = 0 \quad \text{uniformly in } n.$$

We claim that if Assumption 5.1 holds, then (5.7) is true for all $\zeta \in \mathscr{L}^\infty(Q)$ and all $F \in \mathscr{W}(\mathbb{H})$ of the form (5.6) with $\phi = 0$. Note that these F are continuous in $\mathbb{H}$. We first prove that the sequence $\{F(\varrho^n,\mathbf{m}^n)\}$ if equi-integrable, and thus weakly precompact in $\mathscr{L}^1(Q)$. For any $\varepsilon > 0$, there exists $R \geqslant 0$ with

$$\sup_n \int_{Q\cap\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n,\mathbf{m}^n)(t,x)\,dx\,dt \leqslant \frac{\varepsilon}{4L}, \tag{5.8}$$

where $L := \|\varphi\|_{\mathscr{L}^\infty(\mathbb{H})} + |c|$. For all n and all $A \subset Q$ Borel we can then estimate

$$\int_A |F(\varrho^n,\mathbf{m}^n)(t,x)|\,dx\,dt \leqslant \|\varphi\|_{\mathscr{L}^\infty(\mathbb{H})}|A| + L\int_A E(\varrho^n,\mathbf{m}^n)(t,x)\,dx\,dt.$$

We decompose the latter integral and obtain

$$\begin{aligned}\int_A E(\varrho^n,\mathbf{u}^n)(t,x)\,dx\,dt &= \int_{A\cap\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n,\mathbf{u}^n)(t,x)\,dx\,dt \\ &\quad + \int_{A\cap\{E(\varrho^n,\mathbf{m}^n)\leqslant R\}} E(\varrho^n,\mathbf{u}^n)(t,x)\,dx\,dt \leqslant \frac{\varepsilon}{4L} + R|A|\end{aligned}$$

for all n, using (5.8). This implies that

$$\int_A |F(\varrho^n,\mathbf{u}^n)(t,x)|\,dx\,dt \leqslant \frac{\varepsilon}{4} + \Big(\|\varphi\|_{\mathscr{L}^\infty(\mathbb{H})} + LR\Big)|A|$$

for all n, which is less than ε for all $A \subset Q$ with sufficiently small Lebesgue measure. We conclude that the sequence $\{F(\varrho^n,\mathbf{m}^n)\}$ is equi-integrable.

We now define $E_R(\mathbf{z}) := \min\{E(\mathbf{z}), R\}$ and $U_R(\mathbf{z}) := \min\{U(r), R\}$, and set

$$F_R(\mathbf{z}) := \varphi(\mathbf{z})\Big(1 + E_R(\mathbf{z})\Big) + cU_R(\mathbf{z})$$

for all $\mathbf{z} = (r,m) \in \mathbb{H}$ and $R \geqslant 0$. Then we have $F_R \in \mathscr{C}_\mathrm{b}(\mathbb{H})$, and so (5.7) holds for F_R and all $\zeta \in \mathscr{L}^\infty(Q)$. By Assumption 5.1, there exists $R \geqslant 0$ such that

$$\sup_n \int_{Q\cap\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n,\mathbf{m}^n)(t,x)\,dx\,dt \leqslant \frac{\varepsilon}{4L'},$$

where $L' := L\|\zeta\|_{\mathscr{L}^\infty(Q)}$, which implies that

$$\int_Q \int_{\mathbb{H}\cap\{E(\mathbf{z})>R\}} E(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt \leqslant \frac{\varepsilon}{4L'}.$$

For all $\zeta \in \mathscr{L}^\infty(Q)$, we can then estimate

$$\begin{aligned}
&\left| \int_Q \int_{\mathbb{H}} \zeta(t,x) F(\mathbf{z})\, \nu_{(t,x)}(d\mathbf{z})\, dx\, dt - \int_Q \zeta(t,x) F(\varrho^n, \mathbf{m}^n)(t,x)\, dx\, dt \right| \\
&\quad \leqslant \left| \int_Q \int_{\mathbb{H}} \zeta(t,x) F_R(\mathbf{z})\, \nu_{(t,x)}(d\mathbf{z})\, dx\, dt - \int_Q \zeta(t,x) F_R(\varrho^n, \mathbf{m}^n)(t,x)\, dx\, dt \right| \\
&\qquad + L\|\zeta\|_{\mathscr{L}^\infty(Q)} \left(\int_Q \int_{\mathbb{H}\cap\{E(\mathbf{z})>R\}} E(\mathbf{z})\, \nu_{(t,x)}(d\mathbf{z})\, dx\, dt \right) \\
&\qquad + L\|\zeta\|_{\mathscr{L}^\infty(Q)} \left(\sup_n \int_{Q\cap\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n, \mathbf{m}^n)(t,x)\, dx\, dt \right),
\end{aligned}$$

which is less than ε for n large. Therefore the map $\langle\!\langle F(r,m) \rangle\!\rangle$ defined by

$$\langle\!\langle F(r,m) \rangle\!\rangle(t,x) := \int_{\mathbb{H}} F(\mathbf{z})\, \nu_{(t,x)}(d\mathbf{z}) \quad \text{for a.e. } (t,x) \in Q \tag{5.9}$$

is in $\mathscr{L}^1(Q)$, and that

$$F(\varrho^n, \mathbf{m}^n) \rightharpoonup \langle\!\langle F(r,m) \rangle\!\rangle \quad \text{weakly in } \mathscr{L}^1(Q) \tag{5.10}$$

for all $F \in \mathscr{W}(\mathbb{H})$ of the form (5.6) with $\phi = 0$.

For given $\phi \in \mathscr{C}(S^{d-1})$, we now consider the map

$$F_\phi(r,m) := \phi\Big(\frac{m}{|m|}\Big) K(r,m) \quad \text{for all } (r,m) \in \mathbb{H},$$

with kinetic energy defined as $K(0,0) := 0$ and $K(r,m) := \frac{1}{2}|m|^2/r$ otherwise. Note that $K(r,0) = 0$ for all $r \geqslant 0$, so the function F_ϕ is continuous in $(0,\infty) \times \mathbf{R}^d$. It is, however, not continuous at the vacuum $(r,m) = (0,0)$ because K is only lower semicontinuous there. In order to have (5.10) for all $F \in \mathscr{W}(\mathbb{H})$, we will assume that no kinetic energy is concentrated in vacuum, in the following sense:

Assumption 5.2 (Absolute Continuity)**.** For all $0 \leqslant a < b < \infty$ we have

$$\lim_{R\to\infty} \int_{[a,b]\times\mathbf{R}^d} \mathbf{1}_{\{\varrho^n < 1/R\}} K(\varrho^n, \mathbf{m}^n)(t,x)\, dx\, dt = 0 \quad \text{uniformly in } n.$$

Note first that since the kinetic energy can be bounded above by the total energy, we can repeat the argument above to show that the sequence $\{F_\phi(\varrho^n, \mathbf{m}^n)\}$ converges weakly in $\mathscr{L}^1(Q)$ (after extracting another subsequence if necessary). We only need to show that the limit can be represented by the Young measure ν as in (5.10) and (5.9). By Assumption 5.2, there exists $R \geqslant 0$ such that

$$\sup_n \int_{Q\cap\{\varrho^n<1/R\}} K(\varrho^n, \mathbf{m}^n)(t,x)\, dx\, dt \leqslant \frac{\varepsilon}{8L'},$$

where $L := \|\phi\|_{\mathscr{C}(S^{d-1})}$ and $L' := L\|\zeta\|_{\mathscr{L}^\infty(Q)}$, which implies that

$$\int_Q \int_{\mathbb{H}\cap\{r<1/R\}} K(\mathbf{z})\, \nu_{(t,x)}(d\mathbf{z})\, dx\, dt \leqslant \frac{\varepsilon}{8L'}.$$

Recall that $\mathbf{z} = (r, m) \in \mathbb{H}$. By Assumption 5.1, we can choose $R \geqslant 0$ such that

$$\sup_n \int_{Q \cap \{E(\varrho^n, \mathbf{m}^n) > R\}} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt \leqslant \frac{\varepsilon}{8L'},$$
$$\int_Q \int_{\mathbb{H} \cap \{E(\mathbf{z}) > R\}} E(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt \leqslant \frac{\varepsilon}{8L'}.$$

We now choose $\psi \in \mathscr{C}(\mathbb{H}, [0,1])$ with compact support and

$$\psi(\mathbf{z}) = 1 \quad \text{for all } \mathbf{z} \in \mathbb{H} \text{ with } E(\mathbf{z}) \leqslant R \text{ and } r \geqslant 1/R.$$

Let $F_\phi = F_0 + F_1$ with $F_0 := \psi F_\phi \in \mathscr{C}_\mathrm{b}(\mathbb{H})$. For all $\zeta \in \mathscr{L}^\infty(Q)$, we then have

$$\begin{aligned}
&\left| \int_Q \int_{\mathbb{H}} \zeta(t,x) F_\phi(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt - \int_Q \zeta(t,x) F_\phi(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt \right| \\
&\quad \leqslant \left| \int_Q \int_{\mathbb{H}} \zeta(t,x) F_0(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt - \int_Q \zeta(t,x) F_0(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt \right| \\
&\qquad + L\|\zeta\|_{\mathscr{L}^\infty(Q)} \left(\int_Q \int_{\mathbb{H} \cap \{E(\mathbf{z}) > R\}} E(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt \right) \\
&\qquad + L\|\zeta\|_{\mathscr{L}^\infty(Q)} \left(\int_Q \int_{\mathbb{H} \cap \{r < 1/R\}} K(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z})\,dx\,dt \right) \\
&\qquad + L\|\zeta\|_{\mathscr{L}^\infty(Q)} \left(\sup_n \int_{Q \cap \{E(\varrho^n, \mathbf{m}^n) > R\}} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt \right) \\
&\qquad + L\|\zeta\|_{\mathscr{L}^\infty(Q)} \left(\sup_n \int_{Q \cap \{\varrho^n < 1/R\}} K(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt \right),
\end{aligned}$$

which is less than ε if n is large enough, because of (5.7). This proves (5.10) for all functions $F \in \mathscr{W}(\mathbb{H})$, with $\langle\!\langle F(r,m) \rangle\!\rangle \in \mathscr{L}^1(Q)$ defined in (5.9).

Consider now an increasing sequence $\{Q_l\}$ of compact subset $Q_l \subset [0,\infty) \times \mathbf{R}^d$ that converges to the whole space $[0,\infty) \times \mathbf{R}^d$ as $l \to \infty$. Repeating the argument above successively for all l, we obtain the following result:

Proposition 5.3 (Young Measure)**.** *Suppose that the sequence $\{(\varrho^n, \mathbf{m}^n)\}$ of functions generated by the time discretization of Definition 4.4 satisfies*

$$\lim_{R \to \infty} \int_{[a,b] \times \mathbf{R}^d} \mathbf{1}_{\{E(\varrho^n, \mathbf{m}^n) > R\}} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt = 0,$$
$$\lim_{R \to \infty} \int_{[a,b] \times \mathbf{R}^d} \mathbf{1}_{\{\varrho^n < 1/R\}} K(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt = 0$$

uniformly in n, for all $0 \leqslant a < b < \infty$. Then there exist a subsequence (which we still denote by $\{(\varrho^n, \mathbf{m}^n)\}$) and a map $\nu \in \mathscr{L}^\infty_\mathrm{w}([0,\infty) \times \mathbf{R}^d, \mathscr{P}_E(\mathbb{H}))$ with

$$F(\varrho^n, \mathbf{m}^n) \longrightarrow \langle\!\langle F(r,m) \rangle\!\rangle \quad \textit{weakly in } \mathscr{L}^1_\mathrm{loc}([0,\infty) \times \mathbf{R}^d),$$

where $F \in \mathscr{W}(\mathbb{H})$ and $\langle\!\langle F(r,m) \rangle\!\rangle \in \mathscr{L}^1_\mathrm{loc}([0,\infty) \times \mathbf{R}^d)$ is defined by

$$\langle\!\langle F(r,m) \rangle\!\rangle(t,x) := \int_{\mathbb{H}} F(\mathbf{z})\,\nu_{(t,x)}(d\mathbf{z}) \quad \textit{for a.e. } (t,x) \in [0,\infty) \times \mathbf{R}^d.$$

Proposition 5.3 gives sufficient conditions for the absence of concentration effects. Moreover, the Young measure constructed above allows us to represent the limits of nonlinear compositions $\{F(\varrho^n, \mathbf{m}^n)\}$ for all test functions $F \in \mathscr{W}(\mathbb{H})$. Note that after extracting another subsequence if necessary, we may assume that $\langle\!\langle r \rangle\!\rangle = \varrho$ and $\langle\!\langle m \rangle\!\rangle = \mathbf{m}$, with density ϱ and momentum $\mathbf{m}$ given by (5.3).

Proposition 5.4 (Strong Convergence)**.** *Suppose that the sequence $\{(\varrho^n, \mathbf{m}^n)\}$ satisfies the assumptions of Proposition 5.3 and generates a Young measure ν as explained there. Then we have $(\varrho^n, \mathbf{m}^n) \longrightarrow (\varrho, \mathbf{m})$ strongly in $\mathscr{L}^1_{\mathrm{loc}}([0,\infty) \times \mathbf{R}^d)$ for limit functions $(\varrho, \mathbf{m})$ defined in* (5.3)*, if and only if*

$$\int_{[0,\infty)\times\mathbf{R}^d} \zeta(t,x)\ \langle\!\langle F(r,m) \rangle\!\rangle(t,x)\,dx\,dt = \int_{[0,\infty)\times\mathbf{R}^d} \zeta(t,x) F(\varrho,\mathbf{m})(t,x)\,dx\,dt \tag{5.11}$$

for all $\zeta \in \mathscr{D}([0,\infty) \times \mathbf{R}^d)$ and $F \in \mathscr{W}(\mathbb{H})$, with $\langle\!\langle F(r,m) \rangle\!\rangle$ defined in (5.9)*.*

Proof. It is known that property (5.11) of the Young measure is equivalent to

$$(\varrho^n, \mathbf{m}^n) \longrightarrow (\varrho, \mathbf{m}) \quad \text{locally in measure;}$$

see Lemma 5.4.1 in [5]. Together with (5.3) this implies strong convergence. □

6. Measure-Valued Solutions

In this section, we study the consistency of the sequence of functions $\{(\varrho^n, \mathbf{m}^n)\}$ constructed in Definition 4.4, for the system of isentropic Euler equations (1.1). Recall that we have the uniform energy bound

$$\sup_n \operatorname*{ess\,sup}_{t\in[0,\infty)} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx < \infty, \tag{6.1}$$

and the energy dissipation bound

$$\sum_{k=1}^\infty \frac{1}{6}(\tau^n)^2 \int_{\mathbf{R}^d} |\nabla U'(\varrho^n_k)|^2 \varrho^n_k\,dx \leqslant \sum_{k=1}^\infty \bigg(\int_{\mathbf{R}^d} E(\varrho^n_{k-1}, \mathbf{m}^n_{k-1})\,dx - \int_{\mathbf{R}^d} E(\varrho^n_k, \mathbf{m}^n_k)\,dx \bigg), \tag{6.2}$$

The right-hand side of (6.2) is uniformly bounded by the initial total energy. We will also require in the following that the assumptions of Proposition 5.3 are satisfied. That is, for all $0 \leqslant a < b < \infty$ we assume that

$$\lim_{R\to\infty} \int_{[a,b]\times\mathbf{R}^d} \mathbf{1}_{\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt = 0, \tag{6.3}$$

$$\lim_{R\to\infty} \int_{[a,b]\times\mathbf{R}^d} \mathbf{1}_{\{\varrho^n<1/R\}} K(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt = 0, \tag{6.4}$$

uniformly in n. After extracting a subsequence if necessary, the sequence $\{(\varrho^n, \mathbf{m}^n)\}$ generates a Young measure ν as explained in Proposition 5.3. Notice that assumptions (6.3) and (6.4) are quite natural. In particular, they are satisfied whenever the sequences $\{\varrho^n\}$ and $\{\mathbf{u}^n\}$ are uniformly bounded in $\mathscr{L}^\infty$.

We can now state our main result.

Theorem 6.1 (Measure-Valued Solutions)**.** *Let $(\bar{\varrho}, \bar{\mathbf{u}}) \in \mathbb{T}\mathscr{P}(\mathbf{R}^d)$ be given such that $\mathcal{E}(\bar{\varrho}, \bar{\mathbf{m}}) < \infty$ where $\bar{\mathbf{m}} := \bar{\varrho}\bar{\mathbf{u}}$. Suppose that the sequence $\{(\varrho^n, \mathbf{m}^n)\}$ generated by the time discretization of Definition 4.4 for some $\tau^n \to 0$ satisfies*

$$\lim_{R\to\infty} \int_{[a,b]\times\mathbf{R}^d} \mathbf{1}_{\{E(\varrho^n,\mathbf{m}^n)>R\}} E(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt = 0,$$
$$\lim_{R\to\infty} \int_{[a,b]\times\mathbf{R}^d} \mathbf{1}_{\{\varrho^n<1/R\}} K(\varrho^n, \mathbf{m}^n)(t,x)\,dx\,dt = 0$$

uniformly in n, for all $0 \leqslant a < b < \infty$. Then there exist a subsequence (still denoted by $\{(\varrho^n, \mathbf{m}^n)\}$ for simplicity) and a Young measure $\nu \in \mathscr{L}^\infty_{\mathrm{w}}([0,\infty)\times\mathbf{R}^d, \mathscr{P}_E(\mathbb{H}))$ as defined in Proposition 5.3 with the following properties:

(1) **Continuity Equation**
Defining $\varrho := \langle\!\langle r \rangle\!\rangle$ and $\mathbf{m} := \langle\!\langle m \rangle\!\rangle$, we have
$$\partial_t \varrho + \nabla\cdot\mathbf{m} = 0 \tag{6.5}$$
and $\varrho(0,\cdot) = \bar{\varrho}$ in the sense of distributions.

(2) **Modified Momentum Equation**
With the notation above, we have
$$\nabla\cdot\Big(\partial_t\mathbf{m} + \nabla\cdot\langle\!\langle r^{-1} m\otimes m \rangle\!\rangle + \nabla\langle\!\langle P(r)\rangle\!\rangle\Big) = 0 \tag{6.6}$$
and $\mathbf{m}(0,\cdot) = \bar{\mathbf{m}}$ in the sense of distributions.

(3) **Energy Inequality**
For a.e. $0 \leqslant t_1 < t_2 < \infty$ we have the inequality
$$\int_{\mathbf{R}^d} \langle\!\langle E(r,m)\rangle\!\rangle(t_2,x)\,dx \leqslant \int_{\mathbf{R}^d} \langle\!\langle E(r,m)\rangle\!\rangle(t_1,x)\,dx. \tag{6.7}$$
Moreover, we have that $\langle\!\langle E(r,m)\rangle\!\rangle(0,\cdot) = E(\bar{\varrho}, \bar{\mathbf{m}})$ a.e.

We proved the continuity equation and the energy inequality already in [12], so here we will concentrate on the momentum equation. Since we assume initial data with finite energy, the estimate (6.7) and lower semicontinuity imply that the total energy will be finite for all times. Therefore there exists a velocity $\mathbf{u}_t \in \mathscr{L}^2(\mathbf{R}^d, \varrho_t)$ where $\varrho_t := \varrho(t,\cdot)$ such that $\mathbf{m}(t,\cdot) = \varrho_t\mathbf{u}_t$ for a.e. $t \in [0,\infty)$. Then Theorem 8.3.1 in [5] and (6.5) imply that the curve $t \mapsto \varrho_t$ is absolutely continuous with respect to the Wasserstein distance. We refer the reader to [12] for more details.

As explained in Section 1, the momentum equation (1.1) formally follows from (6.6) once we know that the velocity field $\mathbf{u}_t$ is tangent, and thus a gradient vector field. On the level of the time discretization of Definition 4.4 we are working with velocities that are indeed tangent, but in principle this property may be lost in the limit $\tau^n \to 0$. In fact, one can prove the following result: Assume that there exists a sequence of densities ϱ^k that converges weakly in $\mathscr{L}^1(\mathbf{R}^d)$ to a limit ϱ. Consider a sequence of *tangent* vector fields $\mathbf{u}^k \in \mathscr{L}^2(\mathbf{R}^d, \varrho^k)$ with uniformly finite energy such that the momentum $\mathbf{m}^k := \varrho^k\mathbf{u}^k$ converges weakly in $\mathscr{L}^p(\mathbf{R}^d)$ to a limit $\mathbf{m}$, for some $p > 1$. This defines a new velocity $\mathbf{u} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ via $\mathbf{m} =: \varrho\mathbf{u}$, and *any* vector field $\mathbf{u} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ can be approximated in this way, if the densities ϱ^k are allowed to vanish at finer and finer scales; see [13]. This phenomenon is well-known also in the context of variational convergence for Dirichlet forms; see [20]. It has been suggested in [23] to use Young measures to identify necessary conditions for the preservation of the gradient structure for varying densities.

On the other hand, if there exists an open set $\Omega \subset [0,\infty) \times \mathbf{R}^d$ on which the sequence of densities ϱ^n stays uniformly bounded away from zero as $n \to \infty$, then the limit velocity will still be a gradient vector field in Ω. In this case, the Young measure ν constructed in Theorem 6.1 is in fact a gradient Young measure. This concept has been studied in great detail in the literature; see [15].

Proof. We prove (6.6). Let $\zeta \in \mathscr{D}([0,\infty) \times \mathbf{R}^d)$ be given and let $\zeta := \nabla\zeta$. Then

$$\begin{aligned}
&\int_{[0,\infty)\times\mathbf{R}^d} \langle \partial_t \zeta(t,x), \mathbf{m}^n(t,x)\rangle \, dx\, dt + \int_{\mathbf{R}^d} \langle \zeta(0,x), \bar{\mathbf{m}}(x)\rangle \, dx \\
&\quad = \sum_{k=1}^\infty \int_{\mathbf{R}^d} \left\langle \int_{t_{k-1}^n}^{t_k^n} \partial_t \zeta(t,x)\, dt, \mathbf{m}_{k-1}^n(x) \right\rangle dx + \int_{\mathbf{R}^d} \langle \zeta(0,x), \bar{\mathbf{m}}(x)\rangle \, dx \\
&\quad = \sum_{k=1}^\infty \int_{\mathbf{R}^d} \langle \zeta(t_k^n,x), \mathbf{m}_{k-1}^n(x) - \mathbf{m}_k^n(x)\rangle \, dx + \int_{\mathbf{R}^d} \langle \zeta(0,x), \bar{\mathbf{m}}(x) - \mathbf{m}_0^n(x)\rangle \, dx.
\end{aligned}$$

Let us first consider the second integral. Recall that $\bar{\mathbf{m}} = \bar{\varrho}\bar{\mathbf{u}}$ with $(\bar{\varrho}, \bar{\mathbf{u}}) \in \mathbb{T}\mathscr{P}(\mathbf{R}^d)$, and that $\mathbf{m}_0^n = \varrho_0^n \mathbf{u}_0^n$ where $\varrho_0^n = \bar{\varrho}$ and $\mathbf{u}_0^n \in \mathscr{L}^2(\mathbf{R}^d, \varrho_0^n)$ is the velocity projection of Theorem 2.10 corresponding to $\bar{\mathbf{u}}$ and τ^n. By Theorem 3.1, there exists a stress tensor field $\boldsymbol{\sigma}_0^n \in \mathscr{M}(\mathbf{R}^d, \mathcal{S}_+^d)$ with the property that

$$\int_{\mathbf{R}^d} \langle \zeta(0,x), \bar{\mathbf{m}}(x) - \mathbf{m}_0^n(x)\rangle \, dx = -\int_{\mathbf{R}^d} \langle D\zeta(0,x), \boldsymbol{\sigma}_0^n(dx)\rangle,$$

and the latter integral can be estimated as

$$\left| -\int_{\mathbf{R}^d} \langle D\zeta(0,x), \boldsymbol{\sigma}_0^n(dx)\rangle \right| \leqslant \tau^n \|D\zeta(0,\cdot)\|_{\mathscr{L}^\infty(\mathbf{R}^d)} \int_{\mathbf{R}^d} |\bar{\mathbf{u}}(x)|^2 \bar{\varrho}(x)\, dx,$$

which converges to zero as $\tau^n \to 0$. Since $\bar{\mathbf{u}} - \mathbf{u}_0^n \in \mathbb{T}_{\bar{\varrho}}\mathscr{P}(\mathbf{R}^d)$, its $\mathscr{L}^2(\mathbf{R}^d, \bar{\varrho})$-inner product with any divergence-free vector field vanishes. By density, this implies that the limit momentum $\mathbf{m}$ attains the initial data $\bar{\mathbf{m}}$ weakly in $\mathscr{L}^2(\mathbf{R}^d, \bar{\varrho})$.

For all $k \in \mathbf{N}$, let now $\hat{\mathbf{u}}_k^n \in \mathscr{L}^2(\mathbf{R}^d, \varrho_k^n)$ be defined as

$$\int_{\mathbf{R}^d} \varphi(y) \hat{\mathbf{u}}_k^n(y) \varrho_k^n(y)\, dy := \int_{\mathbf{R}^{3d}} \varphi(y) \xi \, \boldsymbol{\alpha}_k^n(d\mathbf{x}, dy)$$

for all $\varphi \in \mathscr{C}_\mathrm{b}(\mathbf{R}^d)$. Thus $\hat{\mathbf{u}}_k^n$ is the velocity transported by the flow. Since

$$(\mathrm{id} + \tau^n \mathbf{u}_{k-1}^n)_\# (\varrho_{k-1}^n \mathcal{L}^d) = \left(\mathrm{id} + \frac{2}{3} (\tau^n)^2 \nabla U'(\varrho_k^n) \right)_\# (\varrho_k^n \mathcal{L}^d),$$

we obtain the following identity: for all $\varphi \in \mathscr{C}_\mathrm{b}(\mathbf{R}^d)$ we have

$$\begin{aligned}
&\int_{\mathbf{R}^d} \varphi(x + \tau^n \mathbf{u}_{k-1}^n(x)) \mathbf{m}_{k-1}^n(x)\, dx \\
&\qquad = \int_{\mathbf{R}^{3d}} \varphi(x + \tau^n \xi) \xi \, \boldsymbol{\alpha}_k^n(d\mathbf{x}, dy) \\
&\qquad = \int_{\mathbf{R}^d} \varphi\left(y + \frac{2}{3} (\tau^n)^2 \nabla U'(\varrho_k^n(y)) \right) \hat{\mathbf{u}}_k^n(y) \varrho_k^n(y)\, dy.
\end{aligned}$$

Writing $\zeta_k^n := \zeta(t_k^n, \cdot)$, we can then decompose

$$\begin{aligned}
&\int_{\mathbf{R}^d} \langle \zeta_k^n, \mathbf{m}_{k-1}^n - \mathbf{m}_k^n \rangle \, dx \\
&\qquad = \int_{\mathbf{R}^d} \langle \zeta_k^n - \zeta_k^n \circ (\mathrm{id} + \tau^n \mathbf{u}_{k-1}^n), \mathbf{m}_{k-1}^n \rangle \, dx \\
&\qquad\quad + \int_{\mathbf{R}^d} \left\langle \zeta_k^n \circ \left(\mathrm{id} + \frac{2}{3} (\tau^n)^2 \nabla U'(\varrho_k^n) \right) - \zeta_k^n, \hat{\mathbf{u}}_k^n \right\rangle \varrho_k^n \, dx \\
&\qquad\quad - \tau^n \int_{\mathbf{R}^d} \langle \zeta_k^n, \nabla U'(\varrho_k^n) \rangle \varrho_k^n \, dx \\
&\qquad\quad + \int_{\mathbf{R}^{2d}} \langle \zeta_k^n(x), \xi - \mathbf{u}_k^n(x) \rangle \boldsymbol{\mu}_k^n(d\mathbf{x});
\end{aligned} \tag{6.8}$$

see (4.11) and the definition of β_τ. We discuss each term in (6.8) individually.

Step 1. To estimate the first term, we define

$$\psi^n(t,x) := \int_0^1 D\zeta(t, x + \theta \tau^n \mathbf{u}_{k-1}^n(x)) \, d\theta \quad \text{for a.e. } x \in \mathbf{R}^d,$$

for all $t \in (t_{k-1}^n, t_k^n]$ and $k \in \mathbf{N}$. Then we can write

$$\int_{\mathbf{R}^d} \langle \zeta_k^n - \zeta_k^n \circ (\mathrm{id} + \tau^n \mathbf{u}_{k-1}^n), \mathbf{m}_{k-1}^n \rangle \, dx = -\tau^n \int_{\mathbf{R}^d} \langle \psi_k^n \mathbf{u}_{k-1}^n, \mathbf{m}_{k-1}^n \rangle \, dx,$$

where $\psi_k^n := \psi^n(t_k^n, \cdot)$. We now use the mean value theorem to estimate

$$\left| \tau^n \psi_k^n(x) - \int_{t_{k-1}^n}^{t_k^n} \psi^n(t,x) \, dt \right| \leqslant (\tau^n)^2 \|\partial_t D\zeta(\cdot, x)\|_{\mathscr{L}^\infty([t_{k-1}^n, t_k^n])} \tag{6.9}$$

for a.e. $x \in \mathbf{R}^d$, which implies that

$$\begin{aligned}
&\left| \int_{\mathbf{R}^d} \langle \zeta_k^n - \zeta_k^n \circ (\mathrm{id} + \tau^n \mathbf{u}_{k-1}^n), \mathbf{m}_{k-1}^n \rangle \, dx + \int_{t_{k-1}^n}^{t_k^n} \int_{\mathbf{R}^d} \langle \psi^n \mathbf{u}^n, \mathbf{m}^n \rangle \, dx \, dt \right| \\
&\qquad \leqslant 2(\tau^n)^2 \|\partial_t D\zeta\|_{\mathscr{L}^\infty([t_{k-1}^n, t_k^n] \times \mathbf{R}^d)} \left(\int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_{k-1}^n|^2 \varrho_{k-1}^n \, dx \right).
\end{aligned}$$

We now sum in k and get

$$\begin{aligned}
&\left| \sum_{k=1}^\infty \int_{\mathbf{R}^d} \langle \zeta_k^n - \zeta_k^n \circ (\mathrm{id} + \tau^n \mathbf{u}_{k-1}^n), \mathbf{m}_{k-1}^n \rangle \, dx + \int_{[0,\infty) \times \mathbf{R}^d} \langle \psi^n \mathbf{u}^n, \mathbf{m}^n \rangle \, dx \, dt \right| \\
&\qquad \leqslant \tau^n \, 2T \|\partial_t D\zeta\|_{\mathscr{L}^\infty([0,\infty) \times \mathbf{R}^d)} \left(\operatorname*{ess\,sup}_{t \in [0,\infty)} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t,x) \, dx \right),
\end{aligned}$$

where $T > 0$ is chosen large enough such that $\operatorname{spt} \zeta \subset [0,T) \times \mathbf{R}^d$. The right-hand side converges to zero as $n \to \infty$ because of the uniform energy bound (6.1).

We now claim that

$$\lim_{n \to \infty} \int_{[0,\infty) \times \mathbf{R}^d} \langle \psi^n \mathbf{u}^n, \mathbf{m}^n \rangle \, dx \, dt = \int_{[0,\infty) \times \mathbf{R}^d} \langle D\zeta, \langle\!\langle r^{-1} m \otimes m \rangle\!\rangle \rangle \, dx \, dt.$$

Indeed, using the mean value theorem, we obtain

$$\left| \psi^n(t,x) - D\zeta(t,x) \right| \leqslant \frac{1}{2} \tau^n |\mathbf{u}_{k-1}^n(x)| \|D^2 \zeta(t, \cdot)\|_{\mathscr{L}^\infty(\mathbf{R}^d)}$$

for a.e. $(t,x) \in (t^n_{k-1}, t^n_k] \times \mathbf{R}^d$ and all $k \in \mathbf{N}$. This implies the estimate

$$\begin{aligned}
&\bigg| \int_{[0,\infty)\times\mathbf{R}^d} \langle \psi^n \mathbf{u}^n, \mathbf{m}^n \rangle \, dx\, dt - \int_{[0,\infty)\times\mathbf{R}^d} \langle (D\zeta)\mathbf{u}^n, \mathbf{m}^n \rangle \, dx\, dt \bigg| \\
&\qquad \leqslant \tau^n \, R \|D^2\zeta\|_{\mathscr{L}^\infty([0,\infty)\times\mathbf{R}^d)} \int_{[0,T]\times\mathbf{R}^d} \mathbf{1}_{\{|\mathbf{u}^n|\leqslant R\}} K(\varrho^n, \mathbf{m}^n) \, dx\, dt \\
&\qquad\quad + 4 \|D\zeta\|_{\mathscr{L}^\infty([0,\infty)\times\mathbf{R}^d)} \int_{[0,T]\times\mathbf{R}^d} \mathbf{1}_{\{|\mathbf{u}^n|>R\}} K(\varrho^n, \mathbf{m}^n) \, dx\, dt,
\end{aligned} \tag{6.10}$$

where $T > 0$ as above. The first integral on the right-hand side is bounded uniformly in n by the energy inequality (6.1). For the second integral, we have

$$\begin{aligned}
\int_{[0,T]\times\mathbf{R}^d} \mathbf{1}_{\{|\mathbf{u}^n|>R\}} K(\varrho^n, \mathbf{m}^n) \, dx\, dt &\leqslant \int_{[0,T]\times\mathbf{R}^d} \mathbf{1}_{\{E(\varrho^n,\mathbf{m}^n)>\frac{1}{2}R\}} E(\varrho^n, \mathbf{m}^n) \, dx\, dt \\
&\quad + \int_{[0,T]\times\mathbf{R}^d} \mathbf{1}_{\{\varrho^n<1/R\}} K(\varrho^n, \mathbf{m}^n) \, dx\, dt.
\end{aligned}$$

Let $\varepsilon > 0$ be given. By choosing $R \geqslant 0$ large enough, we can make the right-hand side less than $\varepsilon/2$ uniformly in n, because of assumptions (6.3) and (6.4). Hence (6.10) converges to zero as $\tau^n \to 0$. On the other hand, we have

$$(\varrho^n)^{-1} \mathbf{m}^n \otimes \mathbf{m}^n \longrightarrow \langle\!\langle r^{-1} m \otimes m \rangle\!\rangle \quad \text{weakly in } \mathscr{L}^1_{\mathrm{loc}}([0,\infty)\times\mathbf{R}^d),$$

by Proposition 5.3, and so our claim follows.

Step 2. To estimate the second term on the right-hand side of (6.8), we write

$$\begin{aligned}
&\int_{\mathbf{R}^d} \Big\langle \zeta^n_k \circ \Big(\mathrm{id} + \frac{2}{3}(\tau^n)^2 \nabla U'(\varrho^n_k) \Big) - \zeta^n_k, \hat{\mathbf{u}}^n_k \Big\rangle \varrho^n_k \, dx \\
&\quad = \frac{2}{3}(\tau^n)^2 \int_{\mathbf{R}^d} \Big\langle \Big(\int_0^1 D\zeta^n_k \circ \Big(\mathrm{id} + \theta \frac{2}{3}(\tau^n)^2 \nabla U'(\varrho^n_k) \Big) \, d\theta \Big) \nabla U'(\varrho^n_k), \hat{\mathbf{u}}^n_k \Big\rangle \varrho^n_k \, dx,
\end{aligned}$$

which implies the estimate

$$\begin{aligned}
&\bigg| \int_{\mathbf{R}^d} \Big\langle \zeta^n_k \circ \Big(\mathrm{id} + \frac{2}{3}(\tau^n)^2 \nabla U'(\varrho^n_k) \Big) - \zeta^n_k, \hat{\mathbf{u}}^n_k \Big\rangle \varrho^n_k \, dx \bigg| \\
&\qquad \leqslant \sqrt{\frac{16}{3}} \tau^n \|D\zeta(t^n_k, \cdot)\|_{\mathscr{L}^\infty(\mathbf{R}^d)} \bigg(\frac{1}{6}(\tau^n)^2 \int_{\mathbf{R}^d} |\nabla U'(\varrho^n_k)|^2 \varrho^n_k \, dx \bigg)^{1/2} \\
&\qquad\quad \times \bigg(\int_{\mathbf{R}^d} \frac{1}{2} |\hat{\mathbf{u}}^n_k|^2 \varrho^n_k \, dx \bigg)^{1/2}.
\end{aligned}$$

Now notice that by definition of $\hat{\mathbf{u}}^n_k$ and Jensen's inequality, we have

$$\int_{\mathbf{R}^d} \frac{1}{2} |\hat{\mathbf{u}}^n_k|^2 \varrho^n_k \, dx \leqslant \int_{\mathbf{R}^{3d}} \frac{1}{2} |\xi|^2 \, \boldsymbol{\alpha}^n_k(d\mathbf{x}, dy) = \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}^n_{k-1}|^2 \varrho^n_{k-1} \, dx, \tag{6.11}$$

which is uniformly bounded by (6.1). We sum in k and obtain

$$\left|\sum_{k=1}^\infty \int_{\mathbf{R}^d} \left\langle \zeta_k^n \circ \left(\mathrm{id} + \frac{2}{3}(\tau^n)^2 \nabla U'(\varrho_k^n)\right) - \zeta_k^n, \hat{\mathbf{u}}_k^n \right\rangle \varrho_k^n \, dx\right| \tag{6.12}$$

$$\leqslant \sqrt{\tau^n} \sqrt{\frac{16T}{3}} \|D\zeta\|_{\mathscr{L}^\infty([0,\infty)\times\mathbf{R}^d)} \left(\sum_{k=1}^\infty \frac{1}{6}(\tau^n)^2 \int_{\mathbf{R}^d} |\nabla U'(\varrho_k^n)|^2 \varrho_k^n \, dx\right)^{1/2}$$

$$\times \left(\operatorname*{ess\,sup}_{t\in[0,\infty)} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t,x)\, dx\right)^{1/2},$$

using the Cauchy-Schwarz inequality. The sum on the right-hand side is controlled by the energy dissipation estimate in (6.2). Hence (6.12) vanishes as $n \to \infty$.

Step 3. Using integration by parts, we can write

$$-\tau^n \int_{\mathbf{R}^d} \langle \zeta_k^n, \nabla U'(\varrho_k^n)\rangle \varrho_k^n \, dx = \tau^n \int_{\mathbf{R}^d} (\nabla \cdot \zeta_k^n)\, P(\varrho_k^n)\, dx.$$

We used the fact that $P(\varrho_k^n) \in \mathscr{W}^{1,1}(\mathbf{R}^d)$ and $\nabla U'(\varrho_k^n)\varrho_k^n = \nabla P(\varrho_k^n)$, which follows from Proposition 4.5 above and Theorem 10.4.6 in [5]. From (6.9), we get

$$\left| - \tau^n \int_{\mathbf{R}^d} \langle \zeta_k^n, \nabla U'(\varrho_k^n)\rangle \varrho_k^n \, dx - \int_{t_{k-1}^n}^{t_k^n} \int_{\mathbf{R}^d} (\nabla\cdot\zeta)\, P(\varrho^n)\, dx\, dt\right|$$

$$\leqslant (\tau^n)^2 \|\partial_t D\zeta\|_{\mathscr{L}^\infty([t_{k-1}^n, t_k^n]\times\mathbf{R}^d)} \left(\int_K P(\varrho_k^n)\, dx\right),$$

where K is a bounded set with $\operatorname{spt}\zeta(t,\cdot) \subset K$ for all $t \in [0,\infty)$. Now we use the estimate $P(r) \leqslant C(1+U(r))$ for all $r\in[0,\infty)$, with $C>0$ some constant. We can then sum in k and arrive at

$$\left|\sum_{k=1}^\infty -\tau^n \int_{\mathbf{R}^d} \langle \zeta_k^n, \nabla U'(\varrho_k^n)\rangle \varrho_k^n\, dx - \int_{[0,\infty)\times\mathbf{R}^d} (\nabla\cdot\zeta)\, P(\varrho^n)\, dx\, dt\right|$$

$$\leqslant \tau^n\, CT \|\partial_t D\zeta\|_{\mathscr{L}^\infty([0,\infty)\times\mathbf{R}^d)} \left(|K| + \operatorname*{ess\,sup}_{t\in[0,\infty)} \int_{\mathbf{R}^d} E(\varrho^n, \mathbf{m}^n)(t,x)\, dx\right).$$

The right-hand side vanishes as $n\to\infty$, and

$$\lim_{n\to\infty} \int_{[0,\infty)\times\mathbf{R}^d} (\nabla\cdot\zeta)\, P(\varrho^n)\, dx\, dt = \int_{[0,\infty)\times\mathbf{R}^d} (\nabla\cdot\zeta)\, \langle\!\langle P(r)\rangle\!\rangle(t,x)\, dx\, dt.$$

Step 4. By Theorem 3.1, there exist $\boldsymbol{\sigma}_k^n \in \mathscr{M}(\mathbf{R}^d, \mathcal{S}_+^d)$ such that

$$\int_{\mathbf{R}^{2d}} \langle \zeta_k^n(x), \xi - \mathbf{u}_k^n(x)\rangle \boldsymbol{\mu}_k^n(d\mathbf{x}) = -\int_{\mathbf{R}^d} \langle D\zeta_k^n(x), \boldsymbol{\sigma}_k^n(dx)\rangle,$$

and the latter integral can be estimated as

$$\left| - \int_{\mathbf{R}^d} \langle D\zeta_k^n(x), \boldsymbol{\sigma}_k^n(dx)\rangle\right| \leqslant 2\tau^n \|D\zeta_k^n\|_{\mathscr{L}^\infty(\mathbf{R}^d)} \left(\int_{\mathbf{R}^d} \frac{1}{2}|\mathbf{u}_{k-1}^n|^2 \varrho_{k-1}^n\, dx\right)^{1/2}$$

$$\times \left(\int_{\mathbf{R}^d} \frac{1}{2}|\mathbf{u}_{k-1}^n|^2 \varrho_{k-1}^n\, dx - \int_{\mathbf{R}^d} \frac{1}{2}|\mathbf{u}_k^n|^2 \varrho_k^n\, dx\right)^{1/2}.$$

We used inequality (6.11). We can now sum in k and obtain

$$
\begin{aligned}
&\left|\sum_{k=1}^{\infty}\int_{\mathbf{R}^{2d}}\langle\zeta_k^n(x),\xi-\mathbf{u}_k^n(x)\rangle\boldsymbol{\mu}_k^n(d\mathbf{x})\right| \qquad (6.13)\\
&\qquad\leqslant\sqrt{\tau^n}\;2\sqrt{T}\|D\zeta\|_{\mathscr{L}^\infty([0,\infty)\times\mathbf{R}^d)}\left(\operatorname*{ess\,sup}_{t\in[0,\infty)}\int_{\mathbf{R}^d}E(\varrho^n,\mathbf{m}^n)(t,x)\,dx\right)^{1/2}\\
&\qquad\quad\times\left(\sum_{k=1}^{\infty}\left(\int_{\mathbf{R}^d}\frac{1}{2}|\mathbf{u}_{k-1}^n|^2\varrho_{k-1}^n\,dx-\int_{\mathbf{R}^d}\frac{1}{2}|\mathbf{u}_k^n|^2\varrho_k^n\,dx\right)\right)^{1/2},
\end{aligned}
$$

using the Cauchy-Schwarz inequality. Hence (6.13) vanishes as $n\to\infty$.

Combining all terms, we find that

$$
\begin{aligned}
&\int_{[0,\infty)\times\mathbf{R}^d}\Big(\langle\partial_t\nabla\varphi,\mathbf{m}\rangle+\langle D^2\varphi,\langle\!\langle r^{-1}m\otimes m\rangle\!\rangle\rangle+\Delta\varphi\;\langle\!\langle P(r)\rangle\!\rangle\Big)\,dx\,dt\\
&\qquad+\int_{\mathbf{R}^d}\langle\nabla\varphi(0,\cdot),\bar{\mathbf{m}}\rangle\,dx=0\quad\text{for all }\varphi\in\mathscr{D}([0,\infty)\times\mathbf{R}^d).
\end{aligned}
$$

This proves the modified momentum equation (6.6) and thus the theorem. □

References

[1] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in* $\mathbf{R}^n$, Math. Z. **230** (1999), no. 2, 259–316.

[2] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.

[3] L. Ambrosio and W. Gangbo, *Hamiltonian ODEs in the Wasserstein space of probability measures*, Comm. Pure Appl. Math. **61** (2007), no. 1, 18–53.

[4] L. Ambrosio and N. Gigli, *Construction of the parallel transport in the Wasserstein space*, Methods Appl. Anal. **15** (2008), no. 1, 1–29.

[5] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, Birkhäuser Verlag, Basel, 2005.

[6] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393.

[7] G. Bouchitté, W. Gangbo, and P. Seppecher, *Michell trusses and lines of principal action*, Math. Models Methods Appl. Sci. **18** (2008), no. 9, 1571–1603.

[8] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417.

[9] C. M. Dafermos, *The entropy rate admissibility criterion for solutions of hyperbolic conservation laws*, J. Differential Equations **14** (1973), 202–212.

[10] W. F. Donoghue, *Distributions and Fourier Transforms*, Pure and Applied Mathematics, vol. 32, Academic Press, New York, 1969.

[11] R. M. Dudley, *On second derivatives of convex functions*, Math. Scand. **41** (1977), no. 1, 159–174.

[12] W. Gangbo and M. Westdickenberg, *Optimal transport for the system of isentropic Euler equations*, Comm. PDE **34** (2009), no. 9, 1041–1073.

[13] N. Gigli, *On the geometry of the space of probability measures endowed with the quadratic optimal transport distance*, Ph.D. Thesis, 2004.

[14] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17.

[15] D. Kinderlehrer and P. Pedregal, *Gradient Young measures generated by sequences in Sobolev spaces*, J. Geom. Anal. **4** (1994), no. 1, 59–90.

[16] P. G. LeFloch and M. Westdickenberg, *Finite energy solutions to the isentropic Euler equations with geometric effects*, J. Math. Pures Appl. (9) **88** (2007), no. 5, 389–429.

[17] P.-L. Lions, *Identification du cône dual des fonctions convexes et applications*, C. R. Acad. Sci. Paris Sér. I Math. **326** (1998), no. 12, 1385–1390.
[18] J. Lott, *Some geometric calculations on Wasserstein space*, Comm. Math. Phys. **277** (2008), no. 2, 423–437.
[19] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179.
[20] U. Mosco, *Composite media and asymptotic Dirichlet forms*, J. Funct. Anal. **123** (1994), no. 2, 368–421.
[21] L. Natile and G. Savaré, *A Wasserstein approach to the one-dimensional sticky particle system*, Preprint (2009).
[22] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101–174.
[23] P. Pedregal, Γ*-convergence through Young measures*, SIAM J. Math. Anal. **36** (2004), no. 2, 423–440.
[24] L. Schwartz, *Théorie des distributions*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X., Hermann, Paris, 1966.
[25] H. Triebel, *Theory of function spaces*, Monographs in Mathematics, vol. 78, Birkhäuser Verlag, Basel, 1983.
[26] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, R.I., 2003.
[27] E. H. Zarantonello, *Projections on convex sets in Hilbert space and spectral theory. I. Projections on convex sets*, Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Madison, WI, 1971), Academic Press, New York, 1971, pp. 237–341.

Michael Westdickenberg, School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0160, U.S.A.
E-mail address: `mwest@math.gatech.edu`