QUARTERLY OF APPLIED MATHEMATICS
VOLUME LXXV, NUMBER 2
JUNE 2017, PAGES 267–285
http://dx.doi.org/10.1090/qam/1459
Article electronically published on November 14, 2016

EULER-LAGRANGE EQUATION FOR A MINIMIZATION PROBLEM OVER MONOTONE TRANSPORT MAPS

By

MICHAEL WESTDICKENBERG

Lehrstuhl für Mathematik (Analysis), RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany

Abstract. A variational time discretization for the compressible Euler equations has been introduced recently. It involves a minimization problem over the cone of monotone transport maps in each timestep. A matrix-valued measure field appears in the minimization as a Lagrange multiplier for the monotonicity constraint. We show that the absolutely continuous part of this measure field vanishes in the support of the density.

Contents

1. Introduction. The compressible Euler equations model the dynamics of compressible fluids like gases. They form a system of hyperbolic conservation laws

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

The unknowns $(\varrho, \boldsymbol{u}, \varepsilon)$ in (1.1) depend on time $t \in [0,\infty)$ and space $x \in \mathbf{R}^d$ and we will assume that suitable initial data $(\varrho, \boldsymbol{u}, \varepsilon)(t = 0, \cdot) =: (\bar\varrho, \bar{\boldsymbol{u}}, \bar\varepsilon)$ is given. We think of ϱ as a map from $[0,\infty)$ into the space of nonnegative, finite Borel measures, which we denote by $\mathscr{M}_+(\mathbf{R}^d)$. The quantity ϱ is called the density and it represents the distribution of

Received March 16, 2016 and, in revised form, October 14, 2016.
2010 *Mathematics Subject Classification.* Primary 35L65, 49J40, 82C40.
Key words and phrases. Compressible gas dynamics, optimal transport.
E-mail address: mwest@instmath.rwth-aachen.de

mass in time and space. The first equation in (1.1) (the continuity equation) expresses the local conservation of mass, where

$$\boldsymbol{u}(t,\cdot) \in \mathscr{L}^2\big(\mathbf{R}^d, \varrho(t,\cdot)\big) \quad \text{for all } t \in [0,\infty) \tag{1.2}$$

is the Eulerian velocity field taking values in $\mathbf{R}^d$. The second equation in (1.1) (the momentum equation) expresses the local conservation of momentum $\boldsymbol{m} := \varrho\boldsymbol{u}$. Note that $\boldsymbol{m}(t,\cdot)$ is a finite $\mathbf{R}^d$-valued Borel measure absolutely continuous with respect to $\varrho(t,\cdot)$ for all $t \in [0,\infty)$, because of (1.2). The quantity ε is the total energy of the fluid and $\varepsilon(t,\cdot)$ is again a measure in $\mathscr{M}_+(\mathbf{R}^d)$ for all $t \in [0,\infty)$. It is reasonable to assume $\varepsilon(t,\cdot)$ to be absolutely continuous with respect to the density $\varrho(t,\cdot)$ (no energy in the vacuum). The third equation in (1.1) (the energy equation) expresses the local conservation of energy. The pressure π is determined by the material and a given function of density and energy. For the case of polytropic gases, the pressure equals $\pi = (\gamma-1)(\varepsilon - \frac{1}{2}\varrho|\boldsymbol{u}|^2)$, with adiabatic coefficient $\gamma > 1$.

As long as solutions are smooth, it is possible to reformulate (1.1) equivalently by substituting for the energy equation a transport equation

$$\partial_t\sigma + \nabla\cdot(\sigma\boldsymbol{u}) = 0 \quad \text{in } [0,\infty)\times\mathbf{R}^d \tag{1.3}$$

for the thermodynamical entropy $\sigma =: \varrho S$, where $S(t,\cdot) \in \mathscr{L}^1(\mathbf{R}^d, \varrho(t,\cdot))$ denotes the specific entropy. The pressure is then given in the form $\pi = \kappa e^S\varrho^\gamma$, with $\kappa > 0$ some constant. In the following, we will utilize this reformulation. It is well known, however, that solutions of hyperbolic conservation laws such as (1.1) typically do not remain smooth globally, but may develop jump discontinuities in finite time, which are called shocks. In this case, the entropy conservation in (1.3) must be relaxed to a differential inequality, as suggested by the second law of thermodynamics.

If the specific entropy S is assumed to be constant in space and time, then (1.1) reduces to the so-called isentropic Euler equations. For smooth solution, the energy equation is implied by the continuity and the momentum equation. For discontinuous solution, however, energy conservation must be relaxed again. The assumption is then that energy can only be dissipated (decreased), but not generated.

Minimizing movements are variational time discretizations generating approximate solutions for evolution equations known as curves of maximal slope; see [14]. They have been studied extensively in recent years in the context of optimal transport theory for certain degenerate parabolic equations; see [3]. Motivated by this research, a variational time discretization for the compressible Euler equations (1.1) has been introduced in [12]. Specifically, assume that for given initial data and timestep $\tau > 0$, the state of the fluid at time $t^k := k\tau$, with $k \in \mathbf{N}$, is approximated by

$$\text{density } \varrho \in \mathscr{P}_2(\mathbf{R}^d), \text{ velocity } \boldsymbol{u} \in \mathscr{L}^2(\mathbf{R}^d, \varrho), \text{ entropy } S \in \mathscr{L}^1(\mathbf{R}^d, \varrho), \tag{1.4}$$

where $\mathscr{P}_2(\mathbf{R}^d)$ is the space of Borel probability measures on $\mathbf{R}^d$ with finite second moment. Without loss of generality, we have assumed that the total mass equals one initially and hence at any positive time. Then density, velocity, and specific entropy at the next time t^{k+1} are determined by a suitable optimization problem, which we interpret (by formal analogy with thermodynamics) as an attempt to *maximize entropy production*: The first law of thermodynamics states that $\Delta U = Q + W$, where ΔU is the change in internal energy, W the work done on the system (nonnegative), and Q is the heat applied to the system. We will aim to maximize the difference $\Delta U - W$. Since at fixed temperature, the change in heat is proportional to the increase in entropy, this amounts to maximizing entropy production.

The work done to the system is defined as a minimal cost: Assume that a material point is located initially at a position $x \in \mathbf{R}^d$ with velocity $\xi \in \mathbf{R}^d$. After the time $\tau > 0$, the material point is at position $z \in \mathbf{R}^d$ with new velocity $\zeta \in \mathbf{R}^d$. As the minimal work needed to effect this change, we consider the infimum

$$\inf\left\{ \int_0^\tau |\ddot{X}(s)|^2\, ds \colon (X,\dot{X})(0) = (x,\xi),\ (X,\dot{X})(\tau) = (z,\zeta) \right\} =: c(x,\xi,z,\zeta).$$

The inf is attained for cubic polynomials, and the minimal cost is

$$c(x,\xi,z,\zeta) = \frac{3}{4\tau^2}|(x+\tau\xi)-z|^2 + \left|\zeta - \left(\xi - \frac{3}{2\tau}\Big((x+\tau\xi)-z\Big)\right)\right|^2. \tag{1.5}$$

If we are free to pick the optimal final velocity ζ, then we can make the second term vanish by choosing $\zeta = \xi - \frac{3}{2\tau}((x+\tau\xi)-z)$. We now define the work as

$$W[\boldsymbol{t}|\varrho,\boldsymbol{u}] := \frac{3}{4\tau^2}\int_{\mathbf{R}^d} \big|\big(x+\tau\boldsymbol{u}(x)\big) - \boldsymbol{t}(x)\big|^2\, \varrho(dx),$$

where $\boldsymbol{t} \in \mathscr{L}^2(\mathbf{R}^d,\varrho)$ determines the transport of the material points, whose initial distribution is given by ϱ. The new velocity is obtained by transporting

$$\boldsymbol{w} := \frac{3}{2}\boldsymbol{v} - \frac{1}{2}\boldsymbol{u}, \quad \text{with} \quad \boldsymbol{v} := \frac{\boldsymbol{t}-\mathrm{id}}{\tau} \tag{1.6}$$

the transport velocity, along with the map $\boldsymbol{t}$. Because of (1.6), we have

$$\frac{3}{4\tau^2}\int_{\mathbf{R}^d} \big|\big(x+\tau\boldsymbol{u}(x)\big) - \boldsymbol{t}(x)\big|^2\, \varrho(dx) = \frac{1}{3}\int_{\mathbf{R}^d} |\boldsymbol{w}(x)-\boldsymbol{u}(x)|^2\, \varrho(dx).$$

In order to compute the change ΔU, we compare the internal energy we would observe if no work was done to the system (which means that all material points travel in the direction of their initial velocity), and the internal energy obtained by pushing ϱ forward under the transport map $\boldsymbol{t}$. We will assume from now on that the density ϱ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^d$ so that $\varrho =: r\mathcal{L}^d$ for some nonnegative $r \in \mathscr{L}^1(\mathbf{R}^d)$. For polytropic gases and with specific entropy S, the internal energy of $\boldsymbol{t}\#\varrho$ can be written (formally) as

$$\mathcal{U}[\boldsymbol{t}\#\varrho, \boldsymbol{t}\#\sigma] := \int_{\mathbf{R}^d} U\big(r(x), S(x)\big)\ \det\big(\nabla\boldsymbol{t}(x)\big)^{1-\gamma}\, dx, \tag{1.7}$$

with internal energy $U(r,S) := \frac{\kappa}{\gamma-1}e^S r^\gamma$ for all $r \geqslant 0$ and $S \in \mathbf{R}$. We have used the change of variables formula and the fact that the entropy $\sigma = \varrho S$ is transported along with the fluid; see (1.3). The internal energy of the freely transported fluid is obtained (formally) by the same formula, with $\mathrm{id} + \tau\boldsymbol{u}$ in place of $\boldsymbol{t}$. Since this energy is a given quantity, maximizing $\Delta U - W$ then amounts to minimizing

$$W[\boldsymbol{t}|\varrho,\boldsymbol{u}] + \mathcal{U}[\boldsymbol{t}\#\varrho, \boldsymbol{t}\#\sigma] \tag{1.8}$$

over a suitable set of transport maps $\boldsymbol{t} \in \mathscr{L}^2(\mathbf{R}^d,\varrho)$. Unfortunately, the functional (1.7) is not coercive, does not control the growth of $\nabla\boldsymbol{t}$ and blows up as $\det(\nabla\boldsymbol{t}) \longrightarrow 0$. In particular, there is no natural function space setting (e.g., in terms of a suitable Sobolev space) in which to search for the transport map $\boldsymbol{t}$ that minimizes (1.8). Moreover, even if the existence of a minimizer can be established, the structure of (1.7) makes it difficult to identify the corresponding Euler-Lagrange equations. We refer the reader to [4] for a related discussion of a similar functional.

We therefore consider two modifications. First, we replace (1.7) by

$$\mathcal{U}[\boldsymbol{t}|\varrho, \mathcal{S}] := \int_{\mathbf{R}^d} U\big(r(x), S(x)\big) \det\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{1-\gamma} dx, \tag{1.9}$$

where the gradient of $\boldsymbol{t}$ is replaced by its deformation

$$\operatorname{def}\big(\boldsymbol{t}(x)\big) := \frac{\nabla \boldsymbol{t}(x) + \nabla \boldsymbol{t}(x)^{\mathrm{T}}}{2}.$$

This renders (1.9) a *convex* functional in $\boldsymbol{t}$. Notice that for small $\tau > 0$ the transport map $\boldsymbol{t}$ is expected to be a perturbation of the identity map id whose derivative is the identity matrix $\mathbb{1}$, which is indeed symmetric (so that $\nabla \mathrm{id} = \operatorname{def}(\mathrm{id})$).

Second, we require that the transport maps $\boldsymbol{t}$ be *monotone*; see Definition 2.3 for the precise statement. Again this can be justified by the fact that the minimizer $\boldsymbol{t}$ is a perturbation of the identity if $\tau > 0$ is small. The deformation $\operatorname{def}(\boldsymbol{t}(x))$ in (1.9) is to be understood to involve only the *absolutely continuous part* of the distributional derivative of $\boldsymbol{t}$, which is a measure since $\boldsymbol{t}$ is of bounded variation locally; see below. In particular, jumps in $\boldsymbol{t}$ (cavitation) do not contribute to $\mathcal{U}[\boldsymbol{t}|\varrho, S]$.

Monotonicity was the crucial ingredient in the recent study of the one-dimensional system of pressureless gas dynamics; see [10, 13, 19]. By rephrasing this system of conservation laws in terms of monotone (optimal) transport maps, one can harness classical results on gradient flows in Hilbert spaces to establish well-posedness und semigroup properties. The monotonicity assumption can be linked to the assumption of sticky particle dynamics (upon collision, particles stick together to form larger compounds), which serves as an entropy condition. Unlike the one-dimensional case, in several space dimensions the composition of monotone maps is typically not a monotone map. Using a Lagrangian reformulation of (1.1) and requiring monotonicity for the global transport maps is not natural, which is why we reset the reference configuration in each timestep of our discretization. Notice that we could have restricted $\boldsymbol{t}$ to the even smaller set of *optimal transport maps* in the sense of optimal transport theory, which (for the Wasserstein distance) are gradients of convex functions (which are monotone). The disadvantage here is that the resulting tangent spaces (i.e., the admissible velocity fields) consist of gradient fields only. In contrast, for monotone maps we recover all vector fields in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$. Finally, the monotonicity of the transport map automatically ensures that *matter does not interpenetrate*. The minimizing map is essentially injective. It was shown in [12] that approximate solutions generated by this time discretization satisfy a crucial energy inequality and converge to a measure-valued solution of (1.1).

Before proceeding, let us fix some notation.

Definition 1.1 (Matrices). In the following, we will denote by $\mathcal{M}^d$ the space of real $(d \times d)$-matrices, and by $\mathcal{M}^d_+$ the space of matrices $M \in \mathcal{M}^d$ such that

$$\langle Mv, v\rangle \geqslant 0 \quad \text{for all } v \in \mathbf{R}^d,$$

where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product in $\mathbf{R}^d$. We will refer to matrices in $\mathcal{M}^d_+$ as positive semidefinite *even when not symmetric*. We will denote by $\mathcal{S}^d$ the space of $M \in \mathcal{M}^d$ with $M^{\mathrm{T}} = M$ (symmetric) and define $\mathcal{S}^d_+ := \mathcal{M}^d_+ \cap \mathcal{S}^d$.

Given data (1.4), we aim to minimize the convex functional

$$\Psi[\boldsymbol{t}|\varrho,\boldsymbol{u},S] := W[\boldsymbol{t}|\varrho,\boldsymbol{u}] + \mathcal{U}[\boldsymbol{t}|\varrho,S] \tag{1.10}$$

over the set $\mathscr{C}_\varrho$ of monotone transport maps (see Definition 2.3), which is a closed convex cone in the Hilbert space $\mathscr{L}^2(\mathbf{R}^d,\varrho)$. Any $\boldsymbol{t} \in \mathscr{C}_\varrho$ can be extended to a maximal monotone map whose domain includes the convex open set $\Omega := \operatorname{int}\operatorname{conv}\operatorname{spt}\varrho$; see Lemma 2.6. A minimizing sequence of transport maps in $\mathscr{C}_\varrho$ for the functional (1.10) can be shown to be uniformly bounded in $\mathrm{BV}_{\mathrm{loc}}(\Omega;\mathcal{M}^d)$. Since the functional (1.9) (with $\nabla\boldsymbol{t}$ representing the absolutely continuous part of the distributional derivative of the maximal monotone extension of $\boldsymbol{t}$) is lower semicontinuous with respect to weak* convergence in $\mathrm{BV}_{\mathrm{loc}}(\Omega;\mathcal{M}^d)$, the existence of a unique minimizer of (1.10) follows, from now on denoted by $\boldsymbol{t} \in \mathscr{C}_\varrho$; see Proposition 5.13 in [12].

In order to derive the Euler-Lagrange equations we consider perturbations of (1.10) around the minimizer $\boldsymbol{t} \in \mathscr{C}_\varrho$. Because of the monotonicity constraint, we may only consider directions in which the perturbed transport map is still in $\mathscr{C}_\varrho$, i.e., we can only consider vector fields in the tangent cone of $\mathscr{C}_\varrho$ at the point $\boldsymbol{t}$. If $\boldsymbol{t}$ is uniformly monotone (see Definition 2.1), then this tangent cone equals $\mathscr{L}^2(\mathbf{R}^d,\varrho)$; see Proposition 4.6 in [12]. Otherwise, all we can conclude is the inequality

$$\begin{aligned}\int_\Omega \langle \boldsymbol{a}(x),\zeta(x)\rangle\,\varrho(dx) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad & (1.11)\\ -\int_\Omega P\big(r(x),S(x)\big)\,\det\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma}\mathrm{tr}\bigg(\mathrm{cof}\,\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}}\nabla\zeta(x)\bigg)\,dx \geqslant 0 &\end{aligned}$$

for sufficiently smooth, monotone vector fields ζ and for $\zeta = \pm\boldsymbol{t}$. Here $\boldsymbol{a} := (\boldsymbol{w}-\boldsymbol{u})/\tau$ is the acceleration and $P(r,S) := U'(r,S)r - U(r,S)$ for all $r,S \geqslant 0$ the pressure function (where $'$ denotes differentiation with respect to r). We denote by cof the cofactor matrix. Choosing $\zeta = \pm\boldsymbol{t}$ in (1.11), we obtain

$$\int_\Omega \langle \boldsymbol{a}(x),\boldsymbol{t}(x)\rangle\,\varrho(dx) - d\int_\Omega P\big(r(x),S(x)\big)\,\det\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{1-\gamma}dx = 0. \tag{1.12}$$

More generally, inequality (1.11) implies (at least formally) that the residual

$$\varrho\boldsymbol{a} + \nabla\cdot\Big(P(r,S)\det\big(\mathrm{def}(\boldsymbol{t})\big)^{-\gamma}\,\mathrm{cof}\,\big(\mathrm{def}(\boldsymbol{t})\big)^{\mathrm{T}}\Big)$$

is an element of the *polar cone* of $\mathscr{C}_\varrho$. We have the following representation:

PROPOSITION 1.2 (Stress tensor). Let $(\varrho,\boldsymbol{u},S)$ be given, with

density $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$, velocity $\boldsymbol{u} \in \mathscr{L}^2(\mathbf{R}^d,\varrho)$, entropy $S \in \mathscr{L}^1(\mathbf{R}^d,\varrho)$

nonnegative. Suppose that $\int_{\mathbf{R}^d} U(r(x),S(x))\,dx < \infty$, where $\varrho =: r\mathcal{L}^d$. For $\tau > 0$, let $\boldsymbol{t} \in \mathscr{C}_\varrho$ be the unique minimizer of (1.10) and let $\boldsymbol{a} \in \mathscr{L}^2(\mathbf{R}^d,\varrho)$ be the acceleration as defined above. Then there exists a matrix field $\mathbf{M} \in \mathscr{M}(\bar{\Omega}; \mathcal{S}^d_+)$ such that

$$\begin{aligned}\int_{\bar{\Omega}} \langle \mathbf{M}(dx), \nabla\zeta(x)\rangle &= \int_\Omega \langle \boldsymbol{a}(x), \zeta(x)\rangle \, \varrho(dx) \\ &\quad - \int_\Omega P\big(r(x), S(x)\big) \, \det\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma} \mathrm{tr}\Big(\mathrm{cof}\,\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}} \nabla\zeta(x)\Big)\, dx\end{aligned} \tag{1.13}$$

for all $\zeta \in \mathscr{C}^1_*(\bar{\Omega}; \mathbf{R}^d)$; see (2.4). Here $\Omega := \operatorname{int}\operatorname{conv}\operatorname{spt}\varrho$.

Remark 1.3. Note that by choosing $\zeta = \mathrm{id}$ in (1.13), we obtain the identity

$$\begin{aligned}\int_{\bar{\Omega}} \mathrm{tr}\big(\mathbf{M}(dx)\big) &= \int_\Omega \langle \boldsymbol{a}(x), x\rangle \, \varrho(dx) \\ &\quad - \int_\Omega P\big(r(x), S(x)\big) \, \det\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma} \mathrm{tr}\Big(\mathrm{cof}\,\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}}\Big)\, dx,\end{aligned}$$

which controls the size of $\mathbf{M}$ since both the stress tensor and the pressure tensor are positive semidefinite and all vector fields in (2.4) are in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$.

We conclude that the Euler-Lagrange equation of the constrained minimization of (1.10) over the closed convex cone $\mathscr{C}_\varrho$ of monotone transport maps is given by

$$\varrho\boldsymbol{a} + \nabla\cdot\Big(P(r,S)\det\big(\mathrm{def}(\boldsymbol{t})\big)^{-\gamma}\,\mathrm{cof}\,\big(\mathrm{def}(\boldsymbol{t})\big)^{\mathrm{T}} + \mathbf{M}\Big) = 0$$

in the distributional sense, where $\mathbf{M}$ is implicitly defined through the transport map $\boldsymbol{t}$. Additional information about the minimizing $\boldsymbol{t}$ can now be gleaned from an investigation into the properties of $\mathbf{M}$, which plays the role of a Lagrange multiplier for the monotonicity constraint (similar to how the pressure plays the role of the Lagrange multiplier for the divergence-free condition in the *incompressible* Euler equations). We expect that $\mathbf{M}$ should vanish wherever the transport map $\boldsymbol{t}$ is strictly monotone; see Definition 2.1. As a first step in this direction we prove here

Proposition 1.4 (Support restriction). With $(\varrho, \boldsymbol{u}, S)$ as in Proposition 1.2 and $\tau > 0$, we consider the minimizer $\boldsymbol{t} \in \mathscr{C}_\varrho$ of (1.10) and the measure $\mathbf{M} \in \mathscr{M}(\bar{\Omega}; \mathcal{S}^d_+)$ defined there, with Lebesgue-Radon-Nikodým decomposition

$$\mathbf{M} =: M\,\mathcal{L}^d + \mathbf{M}^s, \quad \mathbf{M}^s \perp \mathcal{L}^d. \tag{1.14}$$

Here $\Omega := \operatorname{int}\operatorname{conv}\operatorname{spt}\varrho$. Then we have $M(x) = 0$ for a.e. $x \in \operatorname{spt}\varrho$.

Remark 1.5. We observe first that the assumption $\int_{\mathbf{R}^d} U(r(x), S(x))\,dx < \infty$ implies $\varrho \in \mathscr{L}^\gamma(\mathbf{R}^d)$ with $\gamma > 1$. For any $\mathcal{S} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ it follows that

$$\varrho\mathcal{S} \in \mathscr{L}^p(\mathbf{R}^d) \quad \text{with} \quad p := \frac{2\gamma}{\gamma+1} > 1$$

because of Hölder inequality. Now we specialize to the one-dimensional case. Whenever $\varrho\mathcal{S} \in \mathscr{L}^p(\mathbf{R})$, its primitive is a Hölder continuous function, hence absolutely continuous with respect to the Lebesgue measure. This remains true when applied to the acceleration $\boldsymbol{a}$ as above. But the primitive of $\varrho\boldsymbol{a}$ equals the measure

$$\mathbf{H}(dx) := P\big(r(x), S(x)\big)\, \det\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma} \mathrm{cof}\,\Big(\mathrm{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}} dx + \mathbf{M}(dx),$$

as follows from (1.13). As the first component is absolutely continuous with respect to the Lebesgue measure, we obtain that $\mathbf{M}(dx) = M(x)\,dx$ and $\mathbf{M}^s$ vanishes. From Proposition 1.4 we conclude that the measure $\mathbf{M}$ vanishes in spt ϱ, so that

$$\varrho \boldsymbol{a} + \nabla \cdot \Big(P(r,S) \ \det \big(\mathrm{def}(\boldsymbol{t})\big)^{-\gamma} \operatorname{cof} \big(\mathrm{def}(\boldsymbol{t})\big)^{\mathrm{T}} \Big) = 0$$

there distributionally; see (1.13). We observe that if spt ϱ consists of disconnected components, then the primitive of the acceleration $\varrho \boldsymbol{a}$ will be constant in the vacuum in between, where it must coincide with $\mathbf{M}$ since the pressure contribution vanishes where the density does. This suggests that the measure $\mathbf{M}$ plays the role of a "virtual pressure" that transports momentum through the vaccum; see also Remark 4.6 of [12]. For a related discussion in the context of Michell trusses we refer the reader to Remark 5.2 of [9]. While $\mathbf{M}$ does not contribute to the acceleration (since it vanishes in spt ϱ and is piecewise constant outside), it does contribute to the energy balance; see Proposition 4.17 of [12]. We do not know whether in the multidimensional case $\mathbf{M}$ must be absolutely continuous with respect to $\mathcal{L}^d$ as well (in which case $\mathbf{M}$ would again vanish in spt ϱ), or whether there may be singular components supported on lower dimensional sets. This will be investigated in future work.

2. Monotone transport maps. To any $\Gamma \subset \mathbf{R}^d \times \mathbf{R}^d$ we associate a set-valued map $u_\Gamma \colon \mathbf{R}^d \longrightarrow P(\mathbf{R}^d)$ by

$$u_\Gamma(x) := \Big\{ y \in \mathbf{R}^d \colon (x,y) \in \Gamma \Big\} \quad \text{for all } x \in \mathbf{R}^d.$$

Here $P(\mathbf{R}^d)$ is the power set of $\mathbf{R}^d$. For any $u \colon \mathbf{R}^d \longrightarrow P(\mathbf{R}^d)$, we denote by

$$\mathrm{dom}(u) := \Big\{ x \in \mathbf{R}^d \colon u(x) \neq \varnothing \Big\},$$
$$\mathrm{graph}(u) := \Big\{ (x,y) \in \mathbf{R}^d \times \mathbf{R}^d \colon y \in u(x) \Big\}$$

its domain and graph, respectively. A subset $\Gamma \subset \mathbf{R}^d \times \mathbf{R}^d$ is called monotone if

$$\langle x_1 - x_2, y_1 - y_2 \rangle \geqslant 0 \quad \text{for any pair of } (x_i, y_i) \in \Gamma.$$

Such a set is called maximal monotone if for any monotone set $\Gamma' \subset \mathbf{R}^d \times \mathbf{R}^d$ with $\Gamma \subset \Gamma'$ we have that $\Gamma = \Gamma'$. Equivalently, the set Γ is maximal monotone if it is impossible to enlarge Γ without destroying monotonicity. We call a set-valued map u as above (maximal) monotone if the set graph(u) is (maximal) monotone.

DEFINITION 2.1. A subset $\Gamma \subset \mathbf{R}^d \times \mathbf{R}^d$ will be called strictly monotone if

$$\langle x_1 - x_2, y_1 - y_2 \rangle > 0 \quad \text{for any } (x_i, y_i) \in \Gamma.$$

The set Γ will be called uniformly monotone if there exists $\alpha > 0$ with

$$\langle x_1 - x_2, y_1 - y_2 \rangle \geqslant \alpha |x_1 - x_2|^2 \quad \text{for any } (x_i, y_i) \in \Gamma.$$

Analogously, we will talk about strictly and uniformly monotone maps.

REMARK 2.2. By Zorn's lemma, any monotone set (any monotone set-valued map) can be extended to a maximal monotone set (map). Typically, this extension is not unique. A maximal monotone extension can be obtained *constructively* as follows: Let $\Gamma \in \mathbf{R}^d \times \mathbf{R}^d$ be monotone. Then, for all $(x, x^*), (y, y^*) \in \mathbf{R}^d \times \mathbf{R}^d$

(1) define the Fitzpatrick function

$$F_\Gamma(x, x^*) := \sup\Big\{\langle z^*, x\rangle + \langle x^*, z\rangle - \langle z^*, z\rangle \colon (z, z^*) \in \Gamma\Big\};$$

(2) compute its Fenchel conjugate

$$F^*_\Gamma(y^*, y) := \sup\Big\{\langle y^*, x\rangle + \langle x^*, y\rangle - F_\Gamma(x, x^*) \colon (x, x^*) \in \mathbf{R}^d \times \mathbf{R}^d\Big\};$$

(3) compute the proximal average

$$N_\Gamma(x, x^*) := \inf\bigg\{\tfrac{1}{2}F_\Gamma(x_1, x_1^*) + \tfrac{1}{2}F^*_\Gamma(x_2^*, x_2) + \tfrac{1}{8}\|x_1 - x_2\|^2 + \tfrac{1}{8}\|x_1^* - x_2^*\|^2 \colon \\ (x, x^*) = \tfrac{1}{2}(x_1, x_1^*) + \tfrac{1}{2}(x_2, x_2^*)\bigg\}.$$

The function N_Γ is lower semicontinuous, convex, and proper, and the set

$$\bar{\Gamma} := \Big\{(x, x^*) \colon N_\Gamma(x, x^*) = \langle x^*, x\rangle\Big\} \tag{2.1}$$

is a maximal monotone extension of Γ. We refer the reader to [5, 17] for details.

In order to construct a maximal monotone extension of a given monotone function, one can also use the fact that the Cayley transform

$$(x, x^*) \mapsto \frac{1}{\sqrt{2}}(x + x^*, x - x^*), \quad (x, x^*) \in \mathbf{R}^d \times \mathbf{R}^d,$$

(which amounts to a rotation of the coordinate system by $\pi/4$) maps the graphs of monotone functions to 1-Lipschitz functions. By Kirszbraun's theorem, a 1-Lipschitz function can be extended to a 1-Lipschitz function on all of $\mathbf{R}^d$, which by the inverse Cayley transform determines a maximal monotone function; see [1]. The usual proof of Kirszbraun's theorem relies on the axiom of choice. In contrast, the extension procedure outlined above (built on the Fitzpatrick function) is completely constructive. It can provide an alternative proof of Kirszbraun's extension theorem; see [6].

If $\varphi \colon \mathbf{R}^d \longrightarrow \mathbf{R} \cup \{+\infty\}$ is l.s.c. and convex, then its subdifferential $\Gamma := \partial\varphi$ is a maximal monotone map. Its Fitzpatrick function equals $F_\Gamma(x, x^*) = \varphi(x) + \varphi^*(x^*)$ for all $(x, x^*) \in \mathbf{R}^d \times \mathbf{R}^d$. Here φ^* is the Fenchel conjugate of φ. Since

$$F^*_\Gamma(y^*, y) = F_\Gamma(y, y^*) \quad \text{for all } (y, y^*) \in \mathbf{R}^d \times \mathbf{R}^d,$$

the proximal average defined above reduces to $N_\Gamma(x, x^*) = \varphi(x) + \varphi^*(x^*)$. Then the set $\bar{\Gamma}$ given in (2.1) coincides precisely with the subdifferential $\partial\varphi$.

Definition 2.3 (Monotone transport maps). For $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$ and $\boldsymbol{t} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$ taking values in $\mathbf{R}^d$, we define the transport plan $\boldsymbol{\gamma_t} := (\mathrm{id}, \boldsymbol{t})\#\varrho$. Then

$$\mathscr{C}_\varrho := \Big\{\boldsymbol{t} \in \mathscr{L}^2(\mathbf{R}^d; \varrho) \colon \operatorname{spt}\boldsymbol{\gamma_t} \text{ is a monotone subset of } \mathbf{R}^d \times \mathbf{R}^d\Big\}. \tag{2.2}$$

Lemma 2.4 (Closed convex cone). $\mathscr{C}_\varrho$ is a closed convex cone in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$.

Proof. We refer the reader to Lemma 4.2 in [12]. □

Since we are not making any assumptions on $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$, its support may be an arbitrary Borel set. The monotonicity constraint and the extension results enable us to work with objects that are defined on a fixed convex subset of $\mathbf{R}^d$:

DEFINITION 2.5 (Associated maps). Let $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$ be given. For given $\boldsymbol{t} \in \mathscr{C}_\varrho$ we will call u the maximal monotone map associated to $\boldsymbol{t}$ if u is the set-valued map induced by the maximal monotone extension of $\Gamma := \operatorname{spt} \boldsymbol{\gamma_t}$ in Remark 2.2.

Rceall that the domain of a maximal monotone map u satisfies

$$\operatorname{int}\operatorname{conv}\operatorname{dom}(u) \subset \operatorname{dom}(u) \subset \operatorname{conv}\operatorname{dom}(u);$$

see Corollary 1.3 in [1]. Here int and conv denote the interior and the closed convex hull of a set, respectively. As a consequence, we obtain the following result:

LEMMA 2.6 (Support). For $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$ and $\boldsymbol{t} \in \mathscr{C}_\varrho$, the domain of the maximal monotone map u associated to $\boldsymbol{t}$ contains the set $\Omega := \operatorname{int}\operatorname{conv}\operatorname{spt}\varrho$.

Proof. We refer the reader to Lemma 3.4 of [12]. □

REMARK 2.7. We pick the construction of Remark 2.2 purely for definiteness. Any other maximal monotone extension of $\Gamma := \operatorname{spt}\boldsymbol{\gamma_t}$ would also work. We prefer to use the construction based on the Fitzpatrick function because it is similar to what is done in optimal transport theory: Here the optimal transport map is contained in the subdifferential of the Kantorovich potential, which solves the associated dual problem and possesses suitable convexity properties. Since the subdifferential of the Kantorovich function is maximal, it defines a natural extension of the optimal transport map to a larger domain; see [3] and the end of Remark 2.2.

Whenever we speak about the derivative of a monotone transport map $\boldsymbol{t} \in \mathscr{C}_\varrho$ we will refer to the derivative of the maximal monotone map associated to $\boldsymbol{t}$. Notice that a maximal monotone map u is locally bounded in the interior of $\operatorname{dom}(u)$ and locally of bounded variation; see Corollary 1.3(3) and Proposition 5.1 in [1].

REMARK 2.8. For any maximal monotone set-valued function $u\colon \mathbf{R}^d \longrightarrow P(\mathbf{R}^d)$ the image $u(x)$ of any $x \in \mathbf{R}^d$ is closed and convex (possibly empty); see Proposition 1.2 of [1]. Therefore the dimension $\dim u(x)$ is well-defined. The singular sets

$$\Sigma^k(u) := \Big\{x \in \mathbf{R}^d\colon\ \dim u(x) \geqslant k\Big\}, \quad \text{with } k = 1\dots d,$$

are countably $\mathcal{H}^{d-k}$-rectifiable; see Theorem 2.2 of [1] for details. Here $\mathcal{H}^n$ denotes the n-dimensional Hausdorff measure. In particular, the set of points $x \in \operatorname{dom}(u)$ for which $u(x)$ contains more than one point (that is, the set $\Sigma^1(u)$) is negligible with respect to the Lebesgue measure $\mathcal{L}^d$. Outside $\Sigma^1(u)$ the function u is continuous. We denote by $m(u)$ the single-valued map that to $x \in \operatorname{dom}(u)$ assigns the element of minimal norm in $u(x)$. Note that $m(u(x))$ is well-defined for all $x \in \operatorname{dom}(u)$.

Proof of Proposition 1.2. The existence of a stress tensor field $\mathbf{M}$ has already been established in Proposition 5.19 in [12]. Here we only explain the necessary modifications needed to obtain the present result, which is slightly more precise. Let

$$\begin{aligned} G_0(v) := &\int_\Omega \langle \boldsymbol{a}(x), u(x)\rangle\, \varrho(dx) \\ &- \int_\Omega P\big(r(x), S(x)\big)\, \det\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma} \operatorname{tr}\bigg(\operatorname{cof}\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}} v(x)\bigg)\, dx \end{aligned} \tag{2.3}$$

for all $v = \operatorname{def}(u)$ and $u \in \mathscr{C}^1_*(\bar\Omega;\mathbf{R}^d)$, where

$$\mathscr{C}^1_*(\bar\Omega;\mathbf{R}^d) := \{u \in \mathscr{C}^1(\mathbf{R}^d;\mathbf{R}^d)\colon \nabla u \in \mathscr{C}_b(\bar\Omega;\mathcal{M}^d)\}. \tag{2.4}$$

Notice that since the integrands in (2.3) are absolutely continuous with respect to the Lebesgue measure, the integration over the boundary $\bar\Omega \setminus \Omega$ is negligible. The functional G_0 is well-defined: Indeed assume that there exists another $\tilde u \in \mathscr{C}^1_*(\bar\Omega;\mathbf{R}^d)$ with $\operatorname{def}(\tilde u(x)) = v(x)$ for all $x \in \bar\Omega$. We define $\bar u(x) := u(x) - \tilde u(x)$ and

$$\operatorname{rot}\big(\bar u(x)\big) := \frac{\nabla \bar u(x) - \nabla \bar u(x)^{\mathrm{T}}}{2} \quad \text{for all } x \in \bar\Omega.$$

We now observe that $\operatorname{def}(\bar u(x))_{i,j} = 0$ and $\partial_k(\operatorname{rot}(\bar u(x))_{i,j}) = 0$ for all $x \in \bar\Omega$ and indices i,j,k. Since $\nabla \bar u = \operatorname{def}(\bar u) + \operatorname{rot}(\bar u)$ and since $\bar\Omega$ is convex (hence connected), we conclude that $\nabla \bar u$ is a constant matrix-valued function with vanishing symmetric part, and so $\bar u$ is a rigid deformation: There exist an antisymmetric matrix $B \in \mathcal{M}^d$ and $c \in \mathbf{R}^d$ such that $\bar u(x) = Bx + c$ for every $x \in \bar\Omega$. The integrands in (2.3) vanish outside of $\bar\Omega$, therefore the behavior of u and $\tilde u$ outside of $\bar\Omega$ is irrelevant for the evaluation of G_0. We may assume that $\bar u$ is a rigid motion defined on all of $\mathbf{R}^d$ so that $\bar u \in \mathscr{C}^1_*(\bar\Omega;\mathbf{R}^d)$. Notice now that both $\bar u$ and $-\bar u$ are monotone maps. Testing the right-hand side of (2.3) with $\pm\bar u$, we therefore obtain that $G_0(\bar u) = 0$. Since G_0 is linear, we conclude that G_0 is indeed well-defined on

$$L := \{\operatorname{def}(u)\colon u \in \mathscr{C}^1_*(\bar\Omega;\mathbf{R}^d)\},$$

which is a subspace of the space $\mathscr{C}_b(\bar\Omega;\mathcal{S}^d)$ of bounded and continuous functions. The functional is positive in the following sense: for all $v \in L \cap C$, with

$$C := \mathscr{C}_b(\bar\Omega;\mathcal{S}^d_+),$$

we have that $G_0(v) \geqslant 0$. Indeed if $v = \operatorname{def}(u) \in C$, then u must be a monotone map; see Theorem 5.3 in [1]. Therefore $G_0(v)$ defined by the right-hand side of (2.3) is nonnegative. We apply a result by Riedl [20] to conclude that G_0 can be extended to a continuous linear map $G\colon \mathscr{C}_b(\bar\Omega;\mathcal{S}^d) \longrightarrow \mathbf{R}$, which moreover is nonnegative when tested against functions in C; see Proposition 2.2 in [11].

We consider now the Stone-Čech compactification $\beta\bar\Omega$ of $\bar\Omega$, which has the property that every map $u \in \mathscr{C}_b(\bar\Omega;\mathcal{S}^d)$ has a continuous extension in the space $\mathscr{C}(\beta\bar\Omega;\mathcal{S}^d)$ of continuous functions on the compact set $\beta\bar\Omega$. We refer the reader to [15] Section 4.8 for details. By the Riesz representation theorem, there exists a finite Radon measure $\mathbf{M} \in \mathscr{M}(\beta\bar\Omega;\mathcal{S}^d)$ that represents the functional G in the sense that

$$G(v) = \int_{\beta\bar\Omega} \langle v(x), \mathbf{M}(dx)\rangle \quad \text{for all } v \in \mathscr{C}(\beta\bar\Omega;\mathcal{S}^d).$$

Since $G(v) \geqslant 0$ for any $v \in \mathscr{C}_b(\beta\bar\Omega;\mathcal{S}^d_+)$ we conclude that $\mathbf{M}$ takes in fact values in $\mathcal{S}^d_+$. Since G is an extension of G_0, we can test the measure $\mathbf{M}$ against a suitable, compactly supported approximation of $u = \mathrm{id}$ and find that $\mathbf{M}$ does not assign any mass to the boundary $\beta\bar{\Omega} \setminus \bar{\Omega}$. We refer the reader to Remark 4.15 in [12]. $\square$

3. Support restriction. In this section, we will establish Proposition 1.4. We will need an approximation of the transport map $\boldsymbol{t} \in \mathscr{C}_\varrho$ that minimizes (1.10) by Lipschitz continuous, monotone maps. This approximation will be provided by the following result.

LEMMA 3.1 (Approximation of monotone maps). For every $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$ absolutely continuous with respect to the Lebesgue measure and for all $\boldsymbol{t} \in \mathscr{C}_\varrho$ there is a sequence of Lipschitz continuous, monotone maps $\boldsymbol{t}_k$ defined on all of $\mathbf{R}^d$, such that

$$\lim_{k\to\infty} \|\boldsymbol{t}_k - \boldsymbol{t}\|_{\mathscr{L}^2(\mathbf{R}^d,\varrho)} = 0.$$

The same statement remains true with $\boldsymbol{t}_k \in \mathscr{C}^1_*(\mathbf{R}^d;\mathbf{R}^d)$ monotone; see (2.4).

Proof. Let u be the maximal monotone map associated to $\boldsymbol{t}$; see Definition 2.5. The domain of u contains the open convex set $\Omega := \operatorname{int}\operatorname{conv}\operatorname{spt}\varrho$, which has Lipschitz boundary. The map u is single-valued outside of a codimension-one rectifiable, hence Lebesgue negligible set. Moreover, for all $x \in \operatorname{dom}(u)$, the image $u(x)$ is nonempty, closed, and convex. Since u is maximal, the map $\mathrm{id} + u$ is surjective (so is $\mathrm{id} + \varepsilon u$ for all $\varepsilon > 0$); see Proposition 1.2 in [1]. We proceed in five steps.

STEP 1. In order to construct the approximating map $\boldsymbol{t}_k$ we consider the resolvent of u and the Yosida approximation. Their definition and properties are well known, but we include here the relevant arguments for the reader's convenience. For $y_i \in \mathbf{R}^d$, $i = 1..2$, and $\varepsilon > 0$ there exist $x_i \in \mathbf{R}^d$ solving the set-valued equation

$$y_i \in x_i + \varepsilon u(x_i)$$

because $\mathrm{id} + \varepsilon u$ is surjective. We write $y_i =: x_i + \varepsilon v_i$ with $v_i \in u(x_i)$. Then

$$\begin{aligned}|y_i - y_2|^2 &= |x_1 - x_2 + \varepsilon(v_1 - v_2)|^2 \\ &= |x_1 - x_2|^2 + \varepsilon^2|v_1 - v_2|^2 + 2\varepsilon\langle v_1 - v_2, x_1 - x_2\rangle.\end{aligned}$$

The last term on the right-hand side is nonnegative, which implies that

$$|x_1 - x_2| \leqslant |y_1 - y_2| \quad \text{and} \quad |v_1 - v_2| \leqslant \frac{1}{\varepsilon}|y_1 - y_2|. \tag{3.1}$$

Taking $y_1 = y_2$ in the first inequality, we conclude that for any y there exists exactly one x with $y \in x + \varepsilon u(x)$. It follows that both the resolvent map $J_\varepsilon := (\mathrm{id} + \varepsilon u)^{-1}$ and the Yosida approximation $u_\varepsilon := (\mathrm{id} - J_\varepsilon)/\varepsilon$ are single-valued, defined on all of $\mathbf{R}^d$, and Lipschitz continuous with Lipschitz constant 1 and $1/\varepsilon$, respectively.

STEP 2. By the very definition of J_ε and u_ε we have that

$$u_\varepsilon(y) = \frac{y - J_\varepsilon(y)}{\varepsilon} \in u\big(J_\varepsilon(y)\big) \quad \text{for all } y \in \mathbf{R}^d.$$

Therefore, since $y_i = J_\varepsilon(y_i) + \varepsilon u_\varepsilon(y_i)$, we obtain

$$\begin{aligned}&\langle u_\varepsilon(y_1) - u_\varepsilon(y_2), y_1 - y_2\rangle \\ &\quad = \langle u_\varepsilon(y_1) - u_\varepsilon(y_2), J_\varepsilon(y_1) - J_\varepsilon(y_2)\rangle + \varepsilon|u_\varepsilon(y_1) - u_\varepsilon(y_2)|^2 \geqslant 0.\end{aligned}$$

We conclude that u_ε is monotone. Since u_ε is also Lipschitz continuous, single-valued, and defined on $\mathbf{R}^d$, it is in fact maximal monotone; see Corollary 1.4 in [1].

For any $x \in \operatorname{dom}(u)$ let $m(u(x))$ be the unique element of minimal norm in $u(x)$. Recall that $u(x)$ is nonempty, closed, and convex. Then we compute

$$\big|u_\varepsilon(x) - m\big(u(x)\big)\big|^2 = -|u_\varepsilon(x)|^2 + \big|m\big(u(x)\big)\big|^2 - 2\big\langle u_\varepsilon(x), m\big(u(x)\big) - u_\varepsilon(x)\big\rangle.$$

But since u is monotone, $m(u(x)) \in u(x)$, and $u_\varepsilon(x) \in u(J_\varepsilon(x))$, we get

$$\big\langle u_\varepsilon(x), m\big(u(x)\big) - u_\varepsilon(x)\big\rangle = \frac{1}{\varepsilon}\big\langle x - J_\varepsilon(x), m\big(u(x)\big) - u_\varepsilon(x)\big\rangle \geqslant 0.$$

Therefore, we have proved the inequality

$$\big|u_\varepsilon(x) - m\big(u(x)\big)\big|^2 \leqslant \big|m\big(u(x)\big)\big|^2 - |u_\varepsilon(x)|^2 \tag{3.2}$$

for all $x \in \operatorname{dom}(u)$, which implies in particular that $|u_\varepsilon(x)| \leqslant |m(u(x))|$ for such x. Since $|x - J_\varepsilon(x)| = \varepsilon |u_\varepsilon(x)|$, it follows that $\lim\limits_{\varepsilon\to 0} J_\varepsilon(x) = x$ for $x \in \operatorname{dom}(u)$.

Step 3. Since $J_\varepsilon(x) = x - \varepsilon u_\varepsilon(x)$ and $u_\varepsilon(x) \in u(J_\varepsilon(x))$, we observe that $y = u_\varepsilon(x)$ is a solution to the inclusion $y \in u(x - \varepsilon y)$. Conversely, any such solution y equals $u_\varepsilon(x)$. Indeed let $z := x - \varepsilon y$. Then the equation becomes $x \in z + \varepsilon u(z)$, thus

$$z = J_\varepsilon(x) \quad \text{and} \quad y = \frac{x - J_\varepsilon(x)}{\varepsilon} = u_\varepsilon(x);$$

see the argument in Step 1. This fact implies that for $\varepsilon, \sigma > 0$ we have

$$u_{\varepsilon+\sigma}(x) = (u_\varepsilon)_\sigma(x).$$

Indeed $y = u_{\varepsilon+\sigma}(x)$ is a solution to the equation

$$y \in u\big(x - (\varepsilon+\sigma)y\big) = u\big((x - \sigma y) - \varepsilon y\big),$$

hence $y = u_\varepsilon(x - \sigma y)$. By applying the above remark again to the Yosida approximation u_ε, which is maximal monotone, we deduce that $y = (u_\varepsilon)_\sigma(x)$.

Now we use (3.2) with u replaced by u_σ. Since u_σ is single-valued, we get

$$|u_{\varepsilon+\sigma}(x) - u_\sigma(x)|^2 \leqslant |u_\sigma(x)|^2 - |u_{\varepsilon+\sigma}(x)|^2.$$

Therefore the map $\varepsilon \mapsto |u_\varepsilon(x)|^2$ is nonincreasing and bounded above by $|m(u(x))|^2$, thus converges to some nonnegative number α as $\varepsilon \to 0$. This implies that

$$\lim_{\varepsilon,\sigma\to 0} |u_{\varepsilon+\sigma}(x) - u_\varepsilon(x)|^2 \leqslant \alpha - \alpha = 0.$$

For any $\varepsilon_k \to 0$, the sequence $u_{\varepsilon_k}(x)$ is therefore a Cauchy sequence, thus converges to some $v \in \mathbf{R}^d$ as $k \to \infty$. Since $u_{\varepsilon_k}(x)$ belongs to $u(J_{\varepsilon_k}(x))$ and the graph of u is closed (see Proposition 1.2 in [1]), we get $v \in u(x)$. Moreover, we have that

$$|v| = \lim_{k\to\infty} |u_{\varepsilon_k}(x)| \leqslant \big|m\big(u(x)\big)\big|.$$

Since $u(x)$ is closed and convex, the projection of 0 onto $u(x)$ is unique, which forces $v = m(u(x))$. In particular, we obtain the same limit for any sequence $\varepsilon_k \to \infty$. Therefore $u_\varepsilon(x)$ converges to $m(u(x))$ as $\varepsilon \to 0$, for all $x \in \operatorname{dom}(u)$.

STEP 4. We fix a sequence $\varepsilon_k \to 0$ and define $\boldsymbol{t}_k := u_{\varepsilon_k}$, where u_ε is the Yosida approximation of u for $\varepsilon > 0$. Then $\boldsymbol{t}_k$ is defined on $\mathbf{R}^d$, single-valued, monotone, and Lipschitz continuous with Lipschitz constant $1/\varepsilon_k$. We also have that

$$|\boldsymbol{t}_k(x)| \leqslant |\boldsymbol{t}(x)| \quad \text{and} \quad \lim_{k\to\infty} \boldsymbol{t}_k(x) = \boldsymbol{t}(x) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d. \tag{3.3}$$

We used Steps 2 and 3, and the fact that the maximal monotone extension u of $\boldsymbol{t}$ is single-valued $\mathcal{L}^d$-a.e. (see Remark 2.8), hence $m(u(x)) = \boldsymbol{t}(x)$ for ϱ-a.e. $x \in \mathbf{R}^d$. For any $R > 0$ we now define the set $E_R := \{x \in \Omega \colon |\boldsymbol{t}(x)| \leqslant R\}$. Then

$$\begin{aligned} \|\boldsymbol{t}_k(x) - \boldsymbol{t}(x)\|^2_{\mathscr{L}^2(E_R^c,\varrho)} &\leqslant \int_{E_R^c} \Big(|\boldsymbol{t}_k(x)| + |\boldsymbol{t}(x)|\Big)^2 \varrho(dx) \\ &\leqslant 4\int_{E_R^c} |\boldsymbol{t}(x)|^2 \varrho(dx) \longrightarrow 0 \quad \text{as } R \to \infty. \end{aligned}$$

Indeed, by definition of E_R and since $\boldsymbol{t} \in \mathscr{L}^2(\mathbf{R}^d, \varrho)$, we can write

$$\int_{E_R^c} |\boldsymbol{t}(x)|^2 \varrho(dx) = \int_{\mathbf{R}^d} |\boldsymbol{t}(x)|^2 \varrho(dx) - \int_{\mathbf{R}^d} \Big(|\boldsymbol{t}(x)|^2 \wedge R^2\Big) \varrho(dx). \tag{3.4}$$

The integrand $|\boldsymbol{t}(x)|^2 \wedge R^2$ converges monotonically to $|\boldsymbol{t}(x)|^2$ for ϱ-a.e. $x \in \mathbf{R}^d$, thus the right-hand side of (3.4) vanishes for $R \to \infty$, because of monotone convergence. On the other hand, we have $|\boldsymbol{t}_k(x) - \boldsymbol{t}(x)| \leqslant 2R$ for ϱ-a.e. $x \in E_R$. Since constants are contained in $\mathscr{L}^1(\mathbf{R}^d, \varrho)$, we conclude using dominated convergence that

$$\|\boldsymbol{t}_k(x) - \boldsymbol{t}(x)\|^2_{\mathscr{L}^2(E_R,\varrho)} \longrightarrow 0 \quad \text{as } k \to \infty, \text{ for any } R > 0.$$

Therefore Lipschitz continuous maps are dense in $\mathscr{C}_\varrho$ in the $\mathscr{L}^2(\mathbf{R}^d, \varrho)$-norm.

STEP 5. We consider a sequence of $\delta_k > 0$ with $\delta_k = o(\varepsilon_k)$ as $k \to \infty$ and define $\mathcal{S}_k := \boldsymbol{t}_k \star \varphi_{\delta_k}$, where $\boldsymbol{t}_k$ is the Yosida approximation of Step 4 and φ_δ is a standard, nonnegative mollifier with compact support, for $\delta > 0$. Since $\boldsymbol{t}_k$ is defined on $\mathbf{R}^d$ and Lipschitz continuous, we have $\mathcal{S}_k \in \mathscr{C}^1(\mathbf{R}^d; \mathbf{R}^d)$ with bounded derivative. The Lipschitz constant of $\boldsymbol{t}_k$ is $1/\varepsilon_k$ for $k \in \mathbf{N}$, which implies the estimate

$$\begin{aligned} |\mathcal{S}_k(x) - \boldsymbol{t}_k(x)| &= \left| \int_{\mathbf{R}^d} \varphi_{\delta_k}(y)\Big(\boldsymbol{t}_k(x-y) - \boldsymbol{t}_k(x)\Big)\, dy \right| \\ &\leqslant \left(\int_{\mathbf{R}^d} \varphi_1(y)|y|\, dy \right) \frac{\delta_k}{\varepsilon_k} \longrightarrow 0 \quad \text{as } k \to \infty, \end{aligned}$$

for all $x \in \mathbf{R}^d$. Using (3.3), we obtain (with $C > 0$ some constant) that

$$|\mathcal{S}_k(x)| \leqslant |\boldsymbol{t}(x)| + C \quad \text{and} \quad \lim_{k\to\infty} \mathcal{S}_k(x) = \boldsymbol{t}(x) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d.$$

Hence $\mathcal{S}_k \longrightarrow \boldsymbol{t}$ in $\mathscr{L}^2(\mathbf{R}^d, \varrho)$ because of dominated convergence. □

The formal proof of Proposition 1.4 requires multiplying the two matrix-valued measures $\mathbf{M}$ and $D\boldsymbol{t}$, which is not defined rigorously. Recall that monotone maps are locally of bounded variation. We will use a suitable approximation. For a similar argument in the context of convex functions we refer the reader to [18].

LEMMA 3.2 (Lower semicontinuity). For $\varrho \in \mathscr{P}_2(\mathbf{R}^d)$ absolutely continuous with respect to the Lebesgue measure, let $\Omega := \operatorname{int}\operatorname{conv}\operatorname{spt}\varrho$. Fix some $\varphi \in \mathscr{D}(\mathbf{R}^d)$ with $\operatorname{spt}\varphi \subset B_1(0)$, $\varphi(\mathbf{R}^d) \subset [0,1]$, and $\int_{\mathbf{R}^d} \varphi(z)\,dz = 1$, and define $\varphi_\varepsilon(z) := \varepsilon^{-d}\varphi(z/\varepsilon)$ for all $z \in \mathbf{R}^d$ and $\varepsilon > 0$. Let u be the maximal monotone map associated to $\boldsymbol{t} \in \mathscr{C}_\varrho$ and u_ε the corresponding Yosida approximation introduced in Lemma 3.1. Finally, let $\mathbf{H} \in \mathscr{M}(\mathbf{R}^d; \mathcal{S}^d_+)$ be given and set $H_\varepsilon\,\mathcal{L}^d := \varphi_\varepsilon \star \mathbf{H}$. Then

$$\liminf_{\varepsilon\to 0} \int_{\mathbf{R}^d} \operatorname{tr}\big(H_\varepsilon(y)\nabla u_\varepsilon(y)\big)\,dy \geqslant \int_\Omega \operatorname{tr}\big(H(x)\nabla u(x)\big)\,dx. \tag{3.5}$$

Here H and ∇u are defined by the Lebesgue-Radon-Nikodým decompositions

$$\begin{aligned} \mathbf{H} &=: H\,\mathcal{L}^d + \mathbf{H}^s, \quad \mathbf{H}^s \perp \mathcal{L}^d, \\ Du &=: \nabla u\,\mathcal{L}^d + D^s u, \quad D^s u \perp \mathcal{L}^d. \end{aligned} \tag{3.6}$$

Proof. Recall that u_ε is Lipschitz continuous, thus ∇u_ε exists almost everywhere. By assumption, we have $H_\varepsilon \in \mathscr{L}^1(\mathbf{R}^d; \mathcal{S}^d_+)$ and $\nabla u_\varepsilon \in \mathscr{L}^\infty(\mathbf{R}^d; \mathcal{M}^d_+)$. Consequently, the map $y \mapsto \operatorname{tr}(H_\varepsilon(y)\nabla u_\varepsilon(y)) \geqslant 0$ is integrable and the integrals on the left-hand side of (3.5) exist for all $\varepsilon > 0$. Consider the map $f_\delta\colon \operatorname{dom}(u) \longrightarrow \mathbf{R}^d$ defined by $f_\delta(x) := x + \delta m(u(x))$ for all $x \in \operatorname{dom}(u)$ and $\delta > 0$, where $m(u(x))$ is the unique element of minimal norm in $u(x)$. Recall that the maximal monotone map u takes values in the subsets of $\mathbf{R}^d$ and $u(x)$ is nonempty, closed, and convex for all $x \in \operatorname{dom}(u)$, so that $m(u(x))$ exists and is finite. Since $m(u(x)) = \lim_{k\to\infty} u_{\varepsilon_k}(x)$ for any sequence $\varepsilon_k \to 0$ and all $x \in \operatorname{dom}(u)$ (see the proof of Lemma 3.1), the map f_δ (being the pointwise limit of a sequence of Lipschitz continuous, therefore Borel measurable maps) is a Borel measurable map. Consider now

$$\Sigma := \big\{x \in \Omega\colon u \text{ is approximately differentiable at } x\big\}.$$

By Proposition 3.71 in [2], the set Σ is Borel measurable. Moreover, the approximate gradient of u coincides with the absolutely continuous part ∇u in (3.6). Then f_δ is approximately differentiable as well, with approximate gradient $\nabla f_\delta(x) = \mathbb{1} + \delta\nabla u(x)$ for all $x \in \Sigma$. Let $\{U_n\}_{n\in\mathbf{N}}$ be a countable base for the topology of $\mathbf{R}^d$. Then

$$\begin{aligned} \operatorname{graph}(f_\delta|_\Sigma) &= \big(\Sigma \times \mathbf{R}^d\big) \setminus \big\{(x,y)\colon y \neq f_\delta(x)\big\} \\ &= \big(\Sigma \times \mathbf{R}^d\big) \setminus \bigcup_{n\in\mathbf{N}} f_\delta^{-1}(U_n) \times U_n^c \end{aligned}$$

is Borel measurable in $\mathbf{R}^d \times \mathbf{R}^d$; see also Proposition 3.1.21 in [22]. Since $f_\delta(\Sigma)$ is the image of the Borel set $\operatorname{graph}(f_\delta|_\Sigma)$ under an orthogonal projection (which is continuous), it is a Souslin set and hence Lebesgue measurable; see Corollary 1.10.9 in [7]. Since the integrand in (3.5) is nonnegative, we can then estimate

$$\int_{\mathbf{R}^d} \operatorname{tr}\big(H_\varepsilon(y)\nabla u_\varepsilon(y)\big)\,dy \geqslant \int_{f_\delta(\Sigma)} \operatorname{tr}\big(H_\varepsilon(y)\nabla u_\varepsilon(y)\big)\,dy \quad \text{for all } \delta > 0.$$

The map f_δ is single-valued and uniformly monotone, and hence a bijection from Σ onto $f_\delta(\Sigma)$. After a change of variables (see e.g. (5.5.2) in [3]), we obtain

$$\int_{f_\delta(\Sigma)} \operatorname{tr}\big(H_\varepsilon(y)\nabla u_\varepsilon(y)\big)\,dy = \int_\Sigma \operatorname{tr}\Big(H_\varepsilon\big(f_\delta(x)\big)\nabla u_\varepsilon\big(f_\delta(x)\big)\Big) \det\big(\nabla f_\delta(x)\big)\,dx.$$

We now set $\delta = \varepsilon$. For all $x \in \Sigma$, we have $u(x) = \{m(u(x))\}$ and

$$\nabla u_\varepsilon\big(f_\varepsilon(x)\big) \det\big(\nabla f_\varepsilon(x)\big) = \frac{\mathbb{1} - \big(\mathbb{1} + \varepsilon \nabla u(x)\big)^{-1}}{\varepsilon} \det\big(\mathbb{1} + \varepsilon \nabla u(x)\big)$$

(see the proof of Lemma 3.1), which converges to $\nabla u(x)$ as $\varepsilon \to 0$.

We claim that $H_\varepsilon(f_\varepsilon(x)) \longrightarrow H(x)$ as $\varepsilon \to 0$, for a.e. $x \in \Sigma$. The sets

$$E_\varepsilon(x) := B_\varepsilon\big(f_\varepsilon(x)\big) \quad \text{for all } \varepsilon > 0$$

are shrinking nicely to x as $\varepsilon \to 0$ (see Section 7.9 in [21]), which means that

$$E_\varepsilon(x) \subset B_{\beta\varepsilon}(x) \quad \text{and} \quad |E_\varepsilon(x)| \geqslant \alpha |B_{\beta\varepsilon}(x)|$$

for $\varepsilon > 0$, for constants $\beta := 1 + |m(u(x))|$ and $\alpha := 1/\beta^d > 0$. Recall that $m(u(x))$ is finite for all $x \in \Sigma$. By definition of H_ε, we then have

$$\begin{aligned} \big|H_{\varepsilon,ij}\big(f_\varepsilon(x)\big) - H_{ij}(x)\big| &= \left| \int_{\mathbf{R}^d} \varphi_\varepsilon\big(f_\varepsilon(x) - z\big)\Big(\mathbf{H}_{ij}(dz) - H_{ij}(x)\,dz\Big)\right| \\ &\leqslant \omega_d \left(\fint_{E_\varepsilon(x)} |H_{ij}(z) - H_{ij}(x)|\,dz + \fint_{E_\varepsilon(x)} |\mathbf{H}^s_{ij}|(dz)\right) \end{aligned} \tag{3.7}$$

for all $i, j = 1 \dots d$. Here $\fint_A := |A|^{-1} \int_A$ is the average over measurable $A \subset \mathbf{R}^d$ with $|A| > 0$ and ω_d is the volume of the d-dimensional unit ball. Both integrals on the right-hand side of (3.7) converge to zero as $\varepsilon \to 0$, for all $x \in \Sigma$ outside a null set N_{ij}; see Theorems 7.10 and 7.13 in [21]. Since $\bigcup_{i,j} N_{ij}$ is negligible, the claim follows. As all integrands are nonnegative, Fatou's lemma implies

$$\begin{aligned} \liminf_{\varepsilon \to 0} \int_\Sigma \operatorname{tr}\Big(H_\varepsilon\big(f_\varepsilon(x)\big) \nabla u_\varepsilon\big(f_\varepsilon(x)\big)\Big) \det\big(\nabla f_\varepsilon(x)\big)\,dx \\ \geqslant \int_\Sigma \operatorname{tr}\big(H(x)\nabla u(x)\big)\,dx. \end{aligned}$$

We may substitute Ω for Σ because $|\Omega \setminus \Sigma| = 0$; see Theorem 3.2 in [1]. $\square$

Lemma 3.3 (Complementarity). Let A and B be symmetric, positive semidefinite, real $(d \times d)$-matrices such that $\operatorname{tr}(AB) = 0$. Let S_+ be the subspace of $\mathbf{R}^d$ spanned by those eigenvectors of A whose corresponding eigenvalues are strictly positive, and let $S_+^\perp$ be its orthogonal complement. Then there exists an orthonormal basis $\{w_k\}$ of eigenvectors of B and a natural number $n \leqslant d$ such that the eigenvectors w_k with $k \geqslant n$ form a basis of S_+ and the corresponding eigenvalues all vanish.

Proof. We proceed by induction over the space dimension. Since the real matrices A and B are symmetric and positive semidefinite, there exist numbers $\lambda_i, \eta_j \geqslant 0$ and orthonormal systems of eigenvectors $\{v_i\}, \{w_j\}$ such that

$$A = \sum_{i=1}^d \lambda_i v_i \otimes v_i \quad \text{and} \quad B = \sum_{j=1}^d \eta_j w_j \otimes w_j. \tag{3.8}$$

Assume now that there exists an index i such that $\lambda_i > 0$. Without loss of generality, we may assume that $i = d$ and that (rotate the coordinate system if necessary) the eigenvector v_d is the dth standard basis vector in $\mathbf{R}^d$. Since v_d is not the zero vector and since the eigenvectors $\{w_j\}$ form an orthonormal basis of $\mathbf{R}^d$, there exists a least one j such that $v_d \cdot w_j \neq 0$. If there is only one such j, then v_d and w_j are collinear. Now notice that from representation (3.8) we obtain that

$$0 = \operatorname{tr}(AB) = \sum_{i,j=1}^{d} \lambda_i \eta_j (v_i \cdot w_j)^2. \tag{3.9}$$

Each term in the sum is nonnegative, and hence for each j such that $v_d \cdot w_j \neq 0$ we must have that $\eta_j = 0$. Relabeling if necessary, we may assume that there exists a natural number $m \leqslant d$ such that $v_d \cdot w_j \neq 0$ for all $j \geqslant m$. Let S_d be the linear subspace of $\mathbf{R}^d$ generated by the eigenvectors w_j with $j \geqslant m$. Since $\eta_j = 0$ for such j, we conclude that S_d is contained in the kernel of B. We also have $v_d \in S_d$. We can then find an orthonormal basis $\tilde{w}_j$, $j \geqslant m$, of S_d such that $\tilde{w}_d = v_d$. Adding the vectors $\tilde{w}_j := w_j$ with $j < m$ we obtain an orthonormal basis of $\mathbf{R}^d$ and

$$B = \sum_{j=1}^{d} \eta_j \tilde{w}_j \otimes \tilde{w}_j.$$

Recall that $\eta_j = 0$ for all $j \geqslant m$. Since $\tilde{w}_d = v_d$ and by orthogonality, we have

$$v_j \cdot \tilde{w}_d = v_d \cdot \tilde{w}_j = 0 \quad \text{for all } j < d. \tag{3.10}$$

We can then rewrite (3.9) in the form

$$\begin{aligned} 0 &= \sum_{i,j=1}^{d} \lambda_i \eta_j (v_i \cdot \tilde{w}_j)^2 \\ &= \sum_{i,j=1}^{d-1} \lambda_i \eta_j (v_i \cdot \tilde{w}_j)^2 + \sum_{j=1}^{d-1} \lambda_d \eta_j (v_d \cdot \tilde{w}_j)^2 + \sum_{i=1}^{d-1} \lambda_i \eta_d (v_i \cdot \tilde{w}_d)^2 + \lambda_d \eta_d. \end{aligned}$$

The last three terms vanish because of (3.10) and since $\eta_d = 0$. The first sum can be reinterpreted as $\operatorname{tr}(\tilde{A}\tilde{B})$ for symmetric, positive semidefinite, real matrices

$$\tilde{A} := \sum_{i=1}^{d-1} \lambda_i v_i \otimes v_i \quad \text{and} \quad \tilde{B} := \sum_{j=1}^{d-1} \eta_j \tilde{w}_j \otimes \tilde{w}_j,$$

mapping $\mathbf{R}^{d-1} \times \{0\} \equiv \mathbf{R}^{d-1}$ to itself. For these we argue analogously. □

Proof of Proposition 1.4. Using the notation of Proposition 1.2, we define

$$\mathbf{H}(dx) := P\big(r(x), S(x)\big)\ \det\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma} \operatorname{cof}\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}} dx + \mathbf{M}(dx),$$

where $\boldsymbol{t} \in \mathscr{C}_\varrho$ is the unique minimizer of (1.10). Let u_ε be the Yosida approximation of the maximal monotone map u associated to $\boldsymbol{t}$, as discussed in Lemma 3.2, and φ_ε the mollifier defined there. By (1.13) and Lemma 3.2, we have

$$\begin{aligned} \liminf_{\varepsilon \to 0} \int_{\mathbf{R}^d} \big\langle \boldsymbol{a}(x), \varphi_\varepsilon \star u_\varepsilon(x) \big\rangle\, \varrho(dx) &\geqslant \int_\Omega \operatorname{tr}\big(M(x) \nabla u(x)\big)\, dx \\ &+ \int_\Omega P\big(r(x), S(x)\big)\ \det\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{-\gamma} \operatorname{tr}\Big(\operatorname{cof}\Big(\operatorname{def}\big(\boldsymbol{t}(x)\big)\Big)^{\mathrm{T}} \nabla u(x)\Big)\, dx. \end{aligned} \tag{3.11}$$

In the last integral we may replace $\nabla u(x)$ by $\nabla \boldsymbol{t}(x)$; see Remark 2.7. We claim now that the integral on the left-hand side of (3.11) converges to $\int_{\mathbf{R}^d} \langle \boldsymbol{a}(x), \boldsymbol{t}(x)\rangle \, \varrho(dx)$. Let us assume for the moment that this is the case. Using (1.12), we obtain

$$\int_\Omega \operatorname{tr}\big(M(x)\nabla u(x)\big)\, dx = 0. \tag{3.12}$$

Since the integrand in (3.12) is nonnegative a.e. (since both $M(x)$ and $\nabla u(x)$ are positive definite, with $M(x)$ being symmetric as well), we have

$$\operatorname{tr}\big(M(x)\nabla u(x)\big) = 0 \quad \text{for a.e. } x \in \Omega.$$

Again we may replace $\nabla u(x)$ by $\nabla \boldsymbol{t}(x)$ for a.e. $x \in \operatorname{spt} \varrho$; see Remark 2.7. Moreover, we may replace $\nabla \boldsymbol{t}(x)$ by its symmetric part $\operatorname{def}(\boldsymbol{t}(x))$. Then Lemma 3.3 implies that there exists an orthonormal basis of $\mathbf{R}^d$ consisting of joint eigenvectors of $M(x)$ and $\operatorname{def}(\boldsymbol{t}(x))$, with the property that each product of corresponding eigenvalues vanishes. But all eigenvalues of $\operatorname{def}(\boldsymbol{t}(x))$ must be strictly positive: The inf of (1.10) is finite (just use $\boldsymbol{t} = \mathrm{id}$ together with our assumption that $\int_{\mathbf{R}^d} U(r(x), S(x))\, dx < \infty$). Consequently, the internal energy (1.9) must be finite if $\boldsymbol{t} \in \mathscr{C}_\varrho$ is the minimizer of (1.10), which implies that $\det(\operatorname{def}(\boldsymbol{t}(x))) > 0$ for a.e. $x \in \operatorname{spt} \varrho$. We conclude that for such x all eigenvalues of $M(x)$ vanish, and so $M(x)$ is the zero matrix.

It only remains to prove the convergence of the left-hand side of (3.11). Arguing as in Steps 4 and 5 of the proof of Lemma 3.1, we obtain the estimates

$$|u_\varepsilon(x)| \leqslant |\boldsymbol{t}(x)| \quad \text{and} \quad |\varphi_\varepsilon \star u_\varepsilon(x)| \leqslant |\boldsymbol{t}(x)| + C \tag{3.13}$$

for ϱ-a.e. $x \in \mathbf{R}^d$ and all $\varepsilon > 0$, where $C > 0$ is some constant. We define

$$\bar\Omega_r := \{x \in \bar B_r(0) \colon \operatorname{dist}(x, \mathbf{R}^d \setminus \Omega \geqslant 1/r\}$$

with $r > 0$. For any $\delta > 0$ there exists an $r > 0$ such that

$$\begin{aligned} &\bigg| \int_{\mathbf{R}^d \setminus \bar\Omega_{2r}} \langle \boldsymbol{a}(x), \boldsymbol{t}(x)\rangle \, \varrho(dx) - \int_{\mathbf{R}^d \setminus \bar\Omega_{2r}} \langle \boldsymbol{a}(x), \varphi_\varepsilon \star u_\varepsilon(x)\rangle \, \varrho(dx) \bigg| \\ &\qquad \leqslant \int_{\mathbf{R}^d \setminus \bar\Omega_{2r}} |\boldsymbol{a}(x)| \big(2|\boldsymbol{t}(x)| + C\big)\, \varrho(dx) \leqslant \delta/2, \end{aligned}$$

because $|\boldsymbol{a}|(2|\boldsymbol{t}| + C)\varrho \in \mathscr{L}^1(\mathbf{R}^d)$. On the compact set $\bar\Omega_r$ the maximal monotone map u is bounded; see Remark 2.7. Shifting the mollifier, we obtain

$$\begin{aligned} &\bigg| \int_{\bar\Omega_r} \langle \boldsymbol{t}(x), \boldsymbol{a}(x)\rangle \, \varrho(dx) - \int_{\bar\Omega_r} \langle \boldsymbol{a}(x), \varphi_\varepsilon \star u_\varepsilon(x)\rangle \, \varrho(dx) \bigg| \\ &\qquad \leqslant \sup u(\bar\Omega_r)\, \|\varrho \boldsymbol{a} - \varphi_\varepsilon \star (\varrho \boldsymbol{a})\|_{\mathscr{L}^1(\mathbf{R}^d)} + \int_{\bar\Omega_r} |\varphi_\varepsilon \star (\boldsymbol{t} - u_\varepsilon)(x)| |\boldsymbol{a}(x)| \, \varrho(dx). \end{aligned} \tag{3.14}$$

Since $\varrho \boldsymbol{a} \in \mathscr{L}^1(\mathbf{R}^d)$, we find that the first term on the right-hand side converges to zero as $\varepsilon \to 0$. On the other hand, we have $\varphi_\varepsilon \star (T - u_\varepsilon)(x) \longrightarrow 0$ pointwise for a.e. $x \in \bar\Omega_r$

since $u_\varepsilon(x) \longrightarrow \boldsymbol{t}(x)$ by Lemma 3.1, with

$$|\varphi_\varepsilon \star (T - u_\varepsilon)(x)| \leqslant 2 \sup u(\bar{\Omega}_r) + C;$$

see (3.13). It follows that the second term on the right-hand of (3.14) vanishes too, by dominated convergence. This proves the claim. $\square$

References

[1] Giovanni Alberti and Luigi Ambrosio, *A geometrical approach to monotone functions in* $\mathbf{R}^n$, Math. Z. **230** (1999), no. 2, 259–316, DOI 10.1007/PL00004691. MR1676726

[2] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000. MR1857292

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008. MR2401600

[4] Roméo Awi and Wilfrid Gangbo, *A polyconvex integrand; Euler-Lagrange equations and uniqueness of equilibrium*, Arch. Ration. Mech. Anal. **214** (2014), no. 1, 143–182, DOI 10.1007/s00205-014-0754-9. MR3237884

[5] Heinz H. Bauschke and Xianfu Wang, *The kernel average for two convex functions and its application to the extension and representation of monotone operators*, Trans. Amer. Math. Soc. **361** (2009), no. 11, 5947–5965, DOI 10.1090/S0002-9947-09-04698-4. MR2529920

[6] Heinz H. Bauschke and Xianfu Wang, *Firmly nonexpansive and Kirszbraun-Valentine extensions: a constructive approach via monotone operator theory*, Nonlinear analysis and optimization I. Nonlinear analysis, Contemp. Math., vol. 513, Amer. Math. Soc., Providence, RI, 2010, pp. 55–64, DOI 10.1090/conm/513/10075. MR2668238

[7] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007. MR2267655

[8] Jonathan M. Borwein and Adrian S. Lewis, *Convex analysis and nonlinear optimization*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3, Springer-Verlag, New York, 2000. Theory and examples. MR1757448

[9] Guy Bouchitté and Giuseppe Buttazzo, *Characterization of optimal shapes and masses through Monge-Kantorovich equation*, J. Eur. Math. Soc. (JEMS) **3** (2001), no. 2, 139–168, DOI 10.1007/s100970000027. MR1831873

[10] Y. Brenier, W. Gangbo, G. Savaré, and M. Westdickenberg, *Sticky particle dynamics with interactions* (English, with English and French summaries), J. Math. Pures Appl. (9) **99** (2013), no. 5, 577–617, DOI 10.1016/j.matpur.2012.09.013. MR3039208

[11] Fabio Cavalletti and Michael Westdickenberg, *The polar cone of the set of monotone maps*, Proc. Amer. Math. Soc. **143** (2015), no. 2, 781–787, DOI 10.1090/S0002-9939-2014-12332-X. MR3283664

[12] F. Cavalletti, M. Sedjro, and M. Westdickenberg, *A variational time discretization for compressible Euler equations* (2014), available at `arXiv:1411.1012v3`.

[13] Fabio Cavalletti, Marc Sedjro, and Michael Westdickenberg, *A simple proof of global existence for the 1D pressureless gas dynamics equations*, SIAM J. Math. Anal. **47** (2015), no. 1, 66–79, DOI 10.1137/130945296. MR3296602

[14] Ennio De Giorgi, *New problems on minimizing movements*, Boundary value problems for partial differential equations and applications, RMA Res. Notes Appl. Math., vol. 29, Masson, Paris, 1993, pp. 81–98. MR1260440

[15] Gerald B. Folland, *Real analysis*, 2nd ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999. Modern techniques and their applications; A Wiley-Interscience Publication. MR1681462

[16] Wilfrid Gangbo and Michael Westdickenberg, *Optimal transport for the system of isentropic Euler equations*, Comm. Partial Differential Equations **34** (2009), no. 7-9, 1041–1073, DOI 10.1080/03605300902892345. MR2560310

[17] N. Ghoussoub, *A variational theory for monotone vector fields*, J. Fixed Point Theory Appl. **4** (2008), no. 1, 107–135.

[18] Pierre-Louis Lions, *Identification du cône dual des fonctions convexes et applications* (French, with English and French summaries), C. R. Acad. Sci. Paris Sér. I Math. **326** (1998), no. 12, 1385–1390, DOI 10.1016/S0764-4442(98)80397-2. MR1649179

[19] Luca Natile and Giuseppe Savaré, *A Wasserstein approach to the one-dimensional sticky particle system*, SIAM J. Math. Anal. **41** (2009), no. 4, 1340–1365, DOI 10.1137/090750809. MR2540269

[20] John Riedl, *Partially ordered locally convex vector spaces and extensions of positive continuous linear mappings*, Math. Ann. **157** (1964), 95–124. MR0169033

[21] Walter Rudin, *Real and complex analysis*, 2nd ed., McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1974. McGraw-Hill Series in Higher Mathematics. MR0344043

[22] S. M. Srivastava, *A course on Borel sets*, Graduate Texts in Mathematics, vol. 180, Springer-Verlag, New York, 1998. MR1619545