SENSITIVITY OF THE COMPLIANCE AND OF THE WASSERSTEIN DISTANCE WITH RESPECT TO A VARYING SOURCE

GUY BOUCHITTÉ[♯], ILARIA FRAGALÀ[*], ILARIA LUCARDESI[♭]

Abstract. We show that the compliance functional in elasticity is differentiable with respect to horizontal variations of the load term, when the latter is given by a possibly concentrated measure; moreover, we provide an integral representation formula for the derivative as a linear functional of the deformation vector field. The result holds true as well for the p-compliance in the scalar case of conductivity. Then we study the limit problem as $p \to +\infty$, which corresponds to differentiate the Wasserstein distance in optimal mass transportation with respect to horizontal perturbations of the two marginals. Also in this case, we obtain an existence result for the derivative, and we show that it is found by solving a minimization problem over the family of all optimal transport plans. When the latter contains only one element, we prove that the derivative of the p-compliance converges to the derivative of the Wasserstein distance in the limit as $p \to +\infty$.

1. Introduction

In many problems from applied mathematics, arising in different areas such as mechanics, dislocation theory, optimal control, or shape optimization, it is an important issue to study rigorously the dependence of solutions not only on the involved shapes but also on the source terms. In case of distributed load terms, the usual method consists in considering the adjoint of the linearized state equation (see for instance [16]), which requires in general some regularity of the involved load. Therefore this approach becomes difficult to handle in case of singular loads. This can be seen even in simpler situations where the sensitivity analysis does not require computing an adjoint state, for instance in the case of the optimization of the scalar p-compliance, which is defined by

$$\mathcal{C}_p(f) := -\inf \Big\{ \int_\Omega \frac{1}{p} |\nabla u|^p \, dx - \langle f, u \rangle \ : \ u \in W^{1,p}(\Omega) \, , \ u = 0 \text{ on } \Sigma \Big\} \, , \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open Lipschitz domain, the exponent p belongs to $(1, +\infty)$, Σ is a closed subset of $\partial\Omega$, and f is an element of $W^{-1,p'}(\overline{\Omega} \setminus \Sigma)$.
If the load f is perturbed into $f_\varepsilon := f + \varepsilon g$, one can easily compute the derivative with respect to the load as

$$\lim_{\varepsilon \to 0} \frac{\mathcal{C}_p(f_\varepsilon) - \mathcal{C}_p(f)}{\varepsilon} = \langle g, u_p \rangle \, , \tag{1.2}$$

being u_p the unique solution to the minimization problem (1.1).
For convenience of the reader, we have enclosed in the Appendix a short proof of (1.2), which holds whenever $g_\varepsilon := \frac{f_\varepsilon - f}{\varepsilon}$ converges strongly to g in $W^{-1,p'}(\overline{\Omega} \setminus \Sigma)$ (and satisfies the additional condition $\langle g_\varepsilon, 1 \rangle = 0$ in case $\Sigma = \emptyset$).
In particular, we emphasize that formula (1.2) is valid only when the perturbation g belongs to $W^{-1,p'}(\overline{\Omega} \setminus \Sigma)$. Unfortunately, there are interesting situations in which this is not the case. Consider for instance the case when $p > n$ and f is a Dirac mass at a point a. If one wants to study the sensitivity of the compliance with respect to the location of a, one has to take $f_\varepsilon := \delta_{a_\varepsilon}$, with $a_\varepsilon := a + \varepsilon h$. In this case, as $g_\varepsilon = \frac{\delta_{a_\varepsilon} - \delta_a}{\varepsilon}$ converges to the dipole distribution $-\operatorname{div}(h\delta_a)$, one would have to apply (1.2) with $g = -\operatorname{div}(h\delta_a)$. Since such a distribution does not belong to $W^{-1,p'}(\overline{\Omega} \setminus \Sigma)$, the meaningfullness of the duality product $\langle g, u_p \rangle$ requires the C^1 regularity of u_p itself, and thus the smoothness of the source term f.

Date: November 21, 2017.
Key words and phrases. Shape derivatives, sensitivity analysis, elastical compliance, transport theory.

The same kind of difficulty occurs when studying the sensitivity of the Monge distance between two probabilities $f^{\pm}$ on $\mathbb{R}^n$, defined by

$$W_1(f^+,f^-) := \inf\Big\{ \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|\, d\gamma(x,y) \ :\ \gamma\in\Pi(f^+,f^-)\Big\}\,,$$

where $\Pi(f^+,f^-)$ is the set of transport plans from f^+ to f^-, *i.e.* $\Pi(f^+,f^-)=\{\gamma\in\mathcal{P}(\overline{\Omega}\times\overline{\Omega})\,,\ \pi_{1\sharp}\gamma=f^+\,,\ \pi_{2\sharp}\gamma=f^-\}$, being $\pi_{1,2}$ the two projections of $\overline{\Omega}\times\overline{\Omega}$ onto the first and second factor.
As proved in the paper [4], this corresponds to consider the limit as $p\to+\infty$ of problem (1.1) when the balanced measure f equals f^+-f^-. This asymptotic result is a natural consequence of the equality $W_1(f^+,f^-)=\mathcal{C}_\infty(f)$, where $\mathcal{C}_\infty(f)$ is given by the dual problem

$$\mathcal{C}_\infty(f) := \max\Big\{\langle f,u\rangle \ :\ u\in \mathrm{Lip}_1(\Omega)\Big\}\,,$$

being $\mathrm{Lip}_1(\Omega)$ the class of 1-Lipschitz functions on Ω. Then, in the smooth case, formula (1.2) continues to hold for $p=\infty$, with u_p replaced by a Monge potential for f (that is a solution to the dual problem $\mathcal{C}_\infty(f)$), provided the latter is unique. For a proof, we refer for instance to [19, Proposition 7.17]. As the Monge potential is in general merely Lipschitz, and further in some cases it is not unique, this formula cannot be applied in absence of regularity (see the related papers [5, 8]), in particular if the dipole distribution g concentrates on the singular set of the potential.
The aim of this paper is to fill this regularity gap and study the sensitivity of both the compliance and the Wasserstein distance under horizontal variations of possibly singular source measures. More precisely, given an open bounded, connected set $\Omega\subset\mathbb{R}^n$ with Lipschitz boundary, and a (scalar or vector) measure f on $\overline{\Omega}$, we consider the horizontal variations f_ε of f induced by a one parameter family of diffeomorphisms Ψ_ε with initial velocity V. Namely, f_ε are defined through their action on any continuous test function u as

$$\langle f_\varepsilon,u\rangle := \langle f, u\circ\Psi_\varepsilon\rangle\,,\qquad \text{with } \Psi_\varepsilon(x)=x+\varepsilon V(x)\,. \tag{1.3}$$

As admissible deformation fields, we consider functions $V\in C^1(\mathbb{R}^n;\mathbb{R}^n)$ such that $V\cdot\nu\le 0$ on $\partial\Omega$, being ν the unit outer normal to $\partial\Omega$.
This kind of variations is reminiscent of domain derivatives in shape sensitivity analysis (*cf.* the monographs [14, 17, 20]). In particular, our approach is intimately connected to our previous works [9, 10].
Our main goals are:

(i) to perform the first derivative in direction V of the compliance functional $\mathcal{C}(f)$, in the general vector setting of elasticity (for the definition of $\mathcal{C}(f)$ and related assumptions on the bulk energy density, including in particular a growth condition or order $p>n$, see the begining of Section 2):

$$\mathcal{C}'(f,V) := \lim_{\varepsilon\to 0}\frac{\mathcal{C}(f_\varepsilon)-\mathcal{C}(f)}{\varepsilon}\,; \tag{1.4}$$

(ii) to perform the first derivative in direction V of the Wasserstein distance

$$\mathcal{C}_\infty'(f,V) := \lim_{\varepsilon\to 0}\frac{\mathcal{C}_\infty(f_\varepsilon)-\mathcal{C}_\infty(f)}{\varepsilon}\,; \tag{1.5}$$

(iii) to study the limit as $p\to+\infty$ of $\mathcal{C}_p'(f,V)$, where $\mathcal{C}_p(f)$ denotes the scalar compliance introduced in (1.1), with $p>n$ and in the Neumann case when $\Sigma=\emptyset$ and f is a balanced load.

In these directions, our achievements can be summarized respectively as follows:

(i) $\mathcal{C}'(f,V)$ exists and is given by

$$\mathcal{C}'(f,V)=\sup_u\inf_\sigma\int_\Omega A(u,\sigma):DV\,dx=\inf_\sigma\sup_u\int_\Omega A(u,\sigma):DV\,dx$$

where $A(u,\sigma)$ is a suitable tensor depending on an optimal displacement u and an optimal field σ for the dual formulation of the compliance (see Theorem 2.2);

(ii) $\mathcal{C}'_\infty(f,V)$ exists and is given by the following expression (in general not linear in V):

$$\mathcal{C}'_\infty(f,V) = \inf\Big\{ \int_{\overline{\Omega}\times\overline{\Omega}} (V(y)-V(x))\cdot \frac{y-x}{|y-x|}\, d\gamma(x,y) \ : \ \gamma \in \operatorname{Argmin}\mathcal{C}_\infty(f)\Big\}$$

(see Theorem 3.3);

(iii) the convergence

$$\lim_{p\to+\infty} \mathcal{C}'_p(f,V) = \mathcal{C}'_\infty(f,V) \tag{1.6}$$

holds true provided there exists a unique optimal transport plan for $\mathcal{C}_\infty(f)$, yielding in particular the linearity of the map $V \mapsto \mathcal{C}'_\infty(f,V)$ (see Theorem 4.1).

These results are stated and proved respectively in Sections 2, 3, and 4, along with several comments and examples.

Hereafter we fix some notation used throughout the paper.

Notation. We denote by $\mathbb{R}^{n\times n}$ the space of $(n\times n)$ matrices with real entries and by $\mathbb{R}^{n\times n}_{\text{sym}}$ the subset of symmetric ones. We denote by I the identity matrix and by 0 the zero one. Given $A \in \mathbb{R}^{n\times n}$, we set A^T its transpose, $A^{\text{sym}} := (A+A^T)/2$ its symmetric part, and $A^{\text{skew}} := (A-A^T)/2$ its skewsymmetric part. When A is invertible, we denote by A^{-1} and A^{-T} the inverse and the transposition of the inverse of A, respectively. Given two vectors a, b in $\mathbb{R}^n$ and two matrices A and B in $\mathbb{R}^{n\times n}$, we use the standard notation $a\cdot b$ and $A : B$ to denote their Euclidean scalar products, namely, adopting the convention of repeated indices, $a\cdot b = a_i b_i$ and $A : B = A_{ij}B_{ij}$; moreover, we define $a\otimes b$ as the tensor product of a and b, namely the matrix $(a\otimes b)_{ij} := a_i b_j$. We endow $\mathbb{R}^{n\times n}$ with the Euclidean norm, that is $|A| := (A : A)^{1/2}$.

Given a tensor field $A \in C^1(\mathbb{R}^n;\mathbb{R}^{n\times n})$, by $\operatorname{div} A$ we mean its divergence with respect to lines, namely $(\operatorname{div} A)_i := \partial_j A_{ij}$.

Given $1 \le p \le +\infty$ we denote by p' its conjugate exponent, defined as usual by the equality $1/p + 1/p' = 1$.

We denote by $\mathcal{D}(\Omega)$ the space of $\mathcal{C}^\infty$ functions having compact support contained into Ω, by $\operatorname{Lip}(\Omega)$ the space of Lipschitz functions on Ω, and by $\operatorname{Lip}_L(\Omega)$ the subspace of L-Lipschitz ones. We denote by $\mathcal{M}(\overline{\Omega};\mathbb{R}^n)$ the space of vector-valued measures with finite total variation on $\overline{\Omega}$, by $\mathcal{M}^+(\overline{\Omega})$ the space of positive measures with finite total variation on $\overline{\Omega}$. Given $\mu \in \mathcal{M}(\overline{\Omega};\mathbb{R}^n)$, we denote by $|\mu| \in \mathcal{M}^+(\overline{\Omega})$ its total variation measure; given $\nu \in \mathcal{M}(\overline{\Omega})$, we denote by $\nu^\pm \in \mathcal{M}^+(\overline{\Omega})$ its positive and negative parts.

In the integrals, unless otherwise indicated, integration is made with respect to the n-dimensional Lebesgue measure. Furthermore, in all the situations when no confusion may arise, we omit to indicate the integration variable.

Whenever we consider L^p-spaces over Ω and over $\partial\Omega$, they are intended the former with respect to the n-dimensional Lebesgue measure over Ω, and the latter with respect to the $(n-1)$-dimensional Hausdorff measure over $\partial\Omega$.

Given a normed vector space X and its topological dual X^*, the duality product between X^* and X is denoted by $\langle\cdot,\cdot\rangle_{(X^*,X)}$; when no ambiguity may arise, we shall omit the subscript. In particular, if $X = \{u \in W^{1,p}(\Omega) : u = 0 \text{ in } \Sigma\}$, then X^* coincides with the space $W^{-1,p'}(\overline{\Omega}\setminus\Sigma)$ consisting of distributions in $W^{-1,p'}(\mathbb{R}^n)$ with support in $\overline{\Omega}\setminus\Sigma$.

2. Derivative of the compliance

In this section we prove the existence and a representation formula for the first derivative of the elastical compliance defined by

$$\mathcal{C}(f) := -\inf\Big\{ \int_\Omega j(e(u))\,dx - \langle f,u\rangle \ : \ u \in \mathcal{U}\Big\}, \tag{2.1}$$

with respect to horizontal variations of the load, defined as in (1.3).

Here $e(u)$ is the symmetric gradient $e(u) = (Du + Du^T)/2$, the space of admissible displacement $\mathcal{U}$ is given by

$$\mathcal{U} := \Big\{ u \in W^{1,p}(\Omega;\mathbb{R}^n) \ : \ u = 0 \text{ on } \Sigma\Big\},$$

for some fixed exponent $p > n$ and a given portion Σ of $\partial\Omega$ with $\mathcal{H}^{n-1}(\Sigma) > 0$, and the potential $j : \mathbb{R}^{n\times n} \to \mathbb{R}$ satisfies the following assumptions

- j is continuous, convex and such that $j(0) = 0$;

– $j(A) = j(A^{\rm sym})$ for every $A \in \mathbb{R}^{n\times n}$;
– there exist two positive constants C_1 and C_2 such that, for some $p > n$

$$C_1(|A^{\rm sym}|^p - 1) \le j(A) \le C_2(|A^{\rm sym}|^p + 1) \quad \forall A \in \mathbb{R}^{n\times n} . \tag{2.2}$$

Notice that $\langle f, u\rangle$ is meant as $\int_{\overline{\Omega}} u\, df$ and makes sense thanks to the continuous embedding of $W^{1,p}(\Omega;\mathbb{R}^n)$ into $C(\overline{\Omega};\mathbb{R}^n)$. In the next lemma (whose proof is postponed below), we show that, under the above assumptions, the infimum problem which defines $\mathcal{C}(f)$ admits an optimal displacement in the space $\mathcal{U}$; moreover, we give the dual formulation of the compliance, which is the key tool to state our result. Let us observe first that, by our assumptions on j, one has $j(Q) = j(0) = 0$ for every skew symmetric Q, thus the Fenchel conjugate j^* satisfies

$$j^*(P) = +\infty \qquad \text{whenever } P^{\rm skew} \neq 0 . \tag{2.3}$$

(indeed $j^*(P) = \sup_{Q\in\mathbb{R}^{n\times n}}\{P : Q - j(Q)\} \ge \sup_{Q\in\mathbb{R}^{n\times n}_{\rm skew}}\{P : Q\}$).

Lemma 2.1. *The compliance functional $\mathcal{C}(f)$ admits the following dual formulation:*

$$\mathcal{C}(f) = \mathcal{C}^*(f) := \inf_{\sigma\in\mathcal{S}}\Big\{ \int_\Omega j^*(\sigma)\, dx \ : \ -\operatorname{div}\sigma = f \ \text{ in } \mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n)\Big\} , \tag{2.4}$$

where

$$\mathcal{S} := \Big\{\sigma \in L^{p'}(\Omega;\mathbb{R}^{n\times n}) \ : \ \operatorname{div}\sigma \in \mathcal{M}(\overline{\Omega};\mathbb{R}^n)\Big\} .$$

Moreover, the infima in (2.1) *and* (2.4) *are attained, and $(u,\sigma) \in \mathcal{U}\times\mathcal{S}$ are optimal respectively for $\mathcal{C}(f)$ and $\mathcal{C}^*(f)$ if and only it holds*

$$\sigma \in \partial j(Du) \ \mathcal{L}^n\text{-a.e. in } \Omega \qquad \text{and} \qquad -\operatorname{div}\sigma = f \ \text{ in } \mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n) . \tag{2.5}$$

Our main result for the derivative of the compliance reads:

Theorem 2.2. *For any admissible deformation field V such that $V(\Sigma) \subseteq \Sigma$, the first derivative of $\mathcal{C}$ at f in direction V defined according to* (1.4) *exists and is given by*

$$\mathcal{C}'(f,V) = \sup_u \inf_\sigma \int_\Omega A(u,\sigma) : DV\, dx = \inf_\sigma \sup_u \int_\Omega A(u,\sigma) : DV\, dx \tag{2.6}$$

where u varies in the set of optimizers of $\mathcal{C}(f)$, σ varies in the set of optimizers of $\mathcal{C}^*(f)$, and $A(u,\sigma)$ is the tensor field

$$A(u,\sigma) := Du\,\sigma - j(Du)\, I . \tag{2.7}$$

Moreover, there exists a saddle point $(u^\star,\sigma^\star)$ at which the sup-inf and inf-sup above are attained.

Remark 2.3. As already mentioned in the Introduction, the assumption $p > n$ serves the purpose of allowing the presence of loads in $\mathcal{M}(\overline{\Omega};\mathbb{R}^n)$, such as Dirac masses. However, we point out that, as it can be seen by inspection of the proof, such assumption is not necessary for the validity of Theorem 2.2 if it is assumed that the load μ belongs to $W^{-1,p'}(\overline{\Omega}\setminus\Sigma)$. In particular Lemma 2.1 still holds in this case (see for instance [3]).

Remark 2.4. In general, equality (2.6) does not allow to conclude that $V \mapsto \mathcal{C}'(f,V)$ is linear, since *a priori* the saddle point $(u^\star,\sigma^\star)$ depends on V. However, the linearity of the derivative can be asserted if the tensor in (2.7) is uniquely determined, as it occurs, for example, when both $\mathcal{C}(f)$ and $\mathcal{C}^*(f)$ have a unique optimizer, or when $\mathcal{C}(f)$ admits a unique optimizer u and j is Gateaux differentiable except at most at the origin. In the latter case the tensor field in (2.7) only depends on u and reads $A(u) = Du\nabla j(Du) - j(Du)\, I$.

Remark 2.5. As a variant, we can consider the scalar setting (either with $\Sigma \neq \emptyset$, or with $\Sigma = \emptyset$ and f balanced real measure, namely $\langle f, u\rangle = 0$ whenever u is constant). In such setting, Theorem 2.2 continues to hold (with the same proof). The unique difference is the algebraic definition of the tensor A appearing in (2.6), which becomes

$$A(u,\sigma) := \nabla u \otimes \sigma - j(\nabla u) I .$$

In particular, when $j(z) = \frac{1}{p}|z|^p$, denoting the compliance functional by $\mathcal{C}_p(f)$, *i.e.*,

$$\mathcal{C}_p(f) := -\inf\Big\{ \int_\Omega \frac{|\nabla u|^p}{p}\, dx - \langle f, u\rangle\Big\} , \tag{2.8}$$

formula (2.6) gives

$$\mathcal{C}_p'(f,V) = \int_\Omega A(u_p) : DV\,dx\,, \tag{2.9}$$

being u_p the (unique) optimal displacement for $\mathcal{C}_p(f)$, and

$$A(u_p) = |\nabla u_p|^{p-2}(\nabla u_p \otimes \nabla u_p) - \frac{|\nabla u_p|^p}{p} I\,. \tag{2.10}$$

As a consequence of Theorem 2.2, by taking vector fields V with compact support in Ω, we obtain the following optimality condition:

Corollary 2.6. *Assume that the tensor field in* (2.7) *is uniquely determined. Then, denoting it simply by A, we have $\mathcal{C}'(f,V) = 0$ for all $V \in \mathcal{D}(\Omega;\mathbb{R}^n)$ if and only if*

$$\operatorname{div} A = 0 \quad \textit{in } \mathcal{D}'(\Omega;\mathbb{R}^n)\,.$$

Below we give some explicit examples of computation of $\mathcal{C}_p'(f,V)$.

Example 2.7. In dimension $n = 1$, let $\Omega = (-1,1)$, and $f = \delta_a - \mathcal{L}^1 \llcorner (-1,0)$, with $a \in (0,1)$. Let us compute $\mathcal{C}_2'(f,V)$ for a deformation field $V \in C^1_c(0,1)$. The optimal function u for $\mathcal{C}_2(f)$ is easily found by using the Euler-Lagrange equation $-u'' = f$ in $\mathcal{D}'(-1,1)$, and $u'(\pm 1) = 0$. We get

$$u = \begin{cases} \frac{x^2}{2} + x & \text{if } x \in (-1,0) \\ x & \text{if } x \in (0,a) \\ a & \text{if } x \in (a,1)\,, \end{cases} \qquad A(u) = \begin{cases} \frac{(x+1)^2}{2} & \text{if } x \in (-1,0) \\ \frac{1}{2} & \text{if } x \in (0,a) \\ 0 & \text{if } x \in (a,1)\,. \end{cases}$$

Then, according to our formula (2.6), we have

$$\mathcal{C}_2'(f,V) = \int_{-1}^1 A(u)V'\,dx = -\int_{-1}^1 (A(u))'V\,dx = \langle \frac{1}{2}\delta_a, V\rangle = \frac{1}{2}V(a)\,.$$

In this elementary example (in which the compliance can be computed explicitly as $\mathcal{C}_2(f) = \frac{a}{2} + \frac{1}{6}$), we already see that u is not differentiable at the point $x = a$ where f is concentrated, so that formula (1.2) cannot be used (taking $g = -(\delta_a)'$).

Example 2.8. In dimension $n = 2$, let Ω be the unit ball $B_1(0)$, and $f = \delta_0 - \frac{1}{2\pi}\mathcal{H}^1 \llcorner \partial B$. For any $p > 2$, the optimal displacement u for $\mathcal{C}_p(f)$ is characterized by the Euler-Lagrange equation $-\Delta_p u = \delta_0$ in $\mathcal{D}'(B_1(0))$, and $-|\nabla u|^{p-2}\frac{\partial u}{\partial \nu} = \frac{1}{2\pi}$ on $\partial B_1(0)$ (with ν the unit outer normal to $\partial B_1(0)$). Thus we get

$$u_p(x) = c|x|^{2-p'}\,, \qquad A(u_p) = \frac{c(1-p')}{|x|^{p'}}\Big(\frac{x}{|x|} \otimes \frac{x}{|x|} - \frac{1}{p}I\Big)\,,$$

where $p' := p/(p-1)$ is the conjugate exponent of p and $c := -\frac{1}{2-p'}(2\pi)^{-\frac{1}{p-1}}$.
We claim that $\mathcal{C}_p'(f,V) = 0$ for every $V \in \mathcal{D}(B_1(0);\mathbb{R}^2)$. By Corollary 2.6, this amounts to show that $\operatorname{div}(A(u_p)) = 0$ in $\mathcal{D}'(B_1(0);\mathbb{R}^2)$. Using the expression of $A(u_p)$, by direct computation it is easy to check that $\operatorname{div} A(u_p) = 0$ in $\mathbb{R}^2 \setminus \{0\}$. On the other hand, let $\varphi \in \mathcal{D}(K;\mathbb{R}^2)$, where K is a compact neighbourhood of the origin. Denoting by ν the unit outer normal to $K \setminus B_\varepsilon(0)$, we get

$$\begin{aligned} \langle -\operatorname{div} A(u_p), \varphi\rangle &= \int_\Omega A(u_p) : D\varphi\,dx = \lim_{\varepsilon\to 0}\int_{K\setminus B_\varepsilon(0)} A(u_p) : D\varphi\,dx \\ &= \lim_{\varepsilon\to 0}\int_{\partial B_\varepsilon(0)} \varphi\cdot(A(u_p)\nu)\,d\mathcal{H}^1 = \lim_{\varepsilon\to 0} -\frac{c(1-p')}{p'|\varepsilon|^{p'}}\int_{\partial B_\varepsilon(0)} \varphi(x)\cdot\frac{x}{|x|}\,d\mathcal{H}^1 = 0 \end{aligned}$$

where the last equality follows from the fact that $\int_{\partial B_\varepsilon(0)} \varphi(x)\cdot\frac{x}{|x|}\,d\mathcal{H}^1 = O(\varepsilon^2)$, and $p' < 2$. As in the previous example, we notice that the solution u_p is not differentiable on the support of f, namely at the point $x = 0$.

The remaining of this section is devoted to the proof of Theorem 2.2.

We start with some preliminary results: first we recall some classical arguments, then we prove Lemma 2.1 stated above.

Lemma 2.9. *Let Y, Z be Banach spaces. Let $A : Y \to Z$ be a linear operator with dense domain $D(A)$. Let $\Phi : Y \to \mathbb{R} \cup \{+\infty\}$ be convex, and $\Psi : Z \to \mathbb{R} \cup \{+\infty\}$ be convex lower semicontinuous. Assume there exists $u_0 \in D(A)$ such that $\Phi(u_0) < +\infty$ and Ψ is continuous at $A\,u_0$. Let Z^* denote the dual space of Z, A^* the adjoint operator of A, and Φ^*, Ψ^* the Fenchel conjugates of Φ, Ψ. Then*

$$-\inf_{u\in Y}\Big\{\Psi(A\,u)+\Phi(u)\Big\} = \inf_{\sigma\in Z^*}\Big\{\Psi^*(\sigma)+\Phi^*(-A^*\,\sigma)\Big\}\,, \tag{2.11}$$

and the infimum at the right hand side is achieved.
Furthermore, $\overline{u}$ and $\overline{\sigma}$ are optimal for the l.h.s. and the r.h.s. of (2.11) *respectively, if and only if there holds $\overline{\sigma} \in \partial\Psi(A\overline{u})$ and $-A^*\overline{\sigma} \in \partial\Phi(\overline{u})$.*

Proof. See [3, Proposition 14]. □

Lemma 2.10. *Let $q \in (1, +\infty)$, let $h : \mathbb{R}^d \to \mathbb{R}$ be a continuous convex integrand satisfying a q-growth condition from both above and below:*

$$C_1(|z|^q - 1) \le h(z) \le C_2(|z|^q + 1)\,.$$

Then the functional $z \mapsto \int_\Omega h(z)\,dx$ is convex, finite, weakly lower semicontinuous and strongly continuous in $L^q(\Omega;\mathbb{R}^d)$.

Proof. See for instance [15, Corollary 6.51 and Theorem 6.54]. □

Lemma 2.11. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary. Let $p \in (1, +\infty)$. Then there exists a positive constant C_p such that*

$$\|Du\|_{L^p(\Omega;\mathbb{R}^{n\times n})} \le C_p\Big(\|u\|_{L^p(\Omega;\mathbb{R}^n)} + \|e(u)\|_{L^p(\Omega;\mathbb{R}^{n\times n})}\Big) \quad \forall u \in W^{1,p}(\Omega;\mathbb{R}^n)\,.$$

Moreover, if Σ is a subset of $\partial\Omega$ with $\mathcal{H}^{n-1}(\Sigma) > 0$, there exists a positive constant $\widetilde{C}_p$ such that

$$\|Dv\|_{L^p(\Omega;\mathbb{R}^{n\times n})} \le \widetilde{C}_p\|e(v)\|_{L^p(\Omega;\mathbb{R}^{n\times n})} \quad \forall v \in W^{1,p}(\Omega;\mathbb{R}^n) \ \ s.t. \ \ v = 0 \text{ on } \Sigma\,.$$

Proof. See, *e.g.*, [18, Chapter 5]. □

Proof of Lemma 2.1. The existence of at least one optimizer for $\mathcal{C}(f)$ follows from Lemmas 2.10 and 2.11 above: the former ensures the lower semicontinuity of the functional

$$\mathcal{U} \ni u \mapsto \int_\Omega j(e(u))\,dx - \langle f, u\rangle\,,$$

and the latter, together with the compact embedding of $W^{1,p}(\Omega;\mathbb{R}^n)$ into $C(\Omega;\mathbb{R}^n)$, gives its coercivity.
The rest of the statement follows by Lemma 2.9 above, applied with $Y = \mathcal{U}$, $Z = L^p(\Omega;\mathbb{R}^{n\times n}_{\rm sym})$, A the symmetric gradient operator, $\Psi(z) := \int_\Omega j(z)\,dx$, and $\Phi(t) := -\langle f, t\rangle$. The convexity and lower semicontinuity of Ψ are ensured by Lemma 2.10, and the convexity of Φ follows from its linearity. Notice that, for $u \in Y = \mathcal{U}$, we have $\Phi(u) = -\int_{\overline{\Omega}} u\,df$. Therefore, for $\sigma \in Z^* = L^{p'}(\Omega;\mathbb{R}^{n\times n}_{\rm sym})$, the Fenchel conjugate $\Phi^*(-A^*\sigma)$ is finite (and equal to zero) if and only if $-\langle \operatorname{div}\sigma, u\rangle_{(Y^*,Y)} = \langle f, u\rangle_{(Y^*,Y)}$. In turn, since any $u \in \mathcal{U}$ is continuous in $\overline{\Omega}$ and vanishes on Σ, this amounts to require that $-\operatorname{div}\sigma = f$ in $\mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n)$. □

Next we write alternative equivalent formulations for the compliance functionals $\mathcal{C}(f_\varepsilon)$. To this aim, for every $\varepsilon \ge 0$, we define $E_\varepsilon : \mathcal{U} \to \mathbb{R}$ and $H_\varepsilon : \mathcal{S} \to \overline{\mathbb{R}}$ as

$$E_\varepsilon(u) := \int_\Omega j((D\Psi_\varepsilon^{-T}Du)^{\rm sym})|\det D\Psi_\varepsilon|\,dx - \langle f, u\rangle\,, \tag{2.12}$$

$$H_\varepsilon(\sigma) := \begin{cases} \displaystyle\int_\Omega j^*(|\det D\Psi_\varepsilon|^{-1}D\Psi_\varepsilon\sigma)|\det D\Psi_\varepsilon|\,dx & \text{if } -\operatorname{div}\sigma = f \text{ in } \mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n)\,, \\ +\infty & \text{otherwise}\,, \end{cases} \tag{2.13}$$

where Ψ_ε are the diffeomorphisms introduced in (1.3). Notice that, according to this definition, we have that $\mathcal{C}(f) = -\inf E_0$ and $\mathcal{C}^*(f) = \inf H_0$. In particular, in view of Lemma 2.1, we have $-\inf E_0 = \inf H_0$. The same equality is valid also for $\varepsilon > 0$, as we prove in the following lemma.

Lemma 2.12. *For every $\varepsilon > 0$, there holds*

$$\mathcal{C}(f_\varepsilon) = -\inf E_\varepsilon = \inf H_\varepsilon \,. \tag{2.14}$$

Proof. Recalling that $f_\varepsilon = (\Psi_\varepsilon)_\sharp f$, for every $u \in \mathcal{U}$, setting $v := u \circ \Psi_\varepsilon$, we may write

$$\int_\Omega j(e(u))\,dx - \langle f_\varepsilon, u\rangle = \int_\Omega j(e(u))\,dx - \langle f, u\circ\Psi_\varepsilon\rangle \tag{2.15}$$

$$= \int_\Omega j(e(v\circ\Psi_\varepsilon^{-1}))\,dx - \langle f, v\rangle = \int_\Omega j((D\Psi_\varepsilon^{-T}Dv)^{\mathrm{sym}})|\det D\Psi_\varepsilon|\,dx - \langle f, v\rangle\,. \tag{2.16}$$

By the arbitrariness of u, since Ψ_ε is a smooth diffeomorphism of $\overline{\Omega}$ into itself, mapping Σ into itself, we infer that the infimum over $u \in \mathcal{U}$ of the left-hand side of (2.15) agrees with the infimum over $v \in \mathcal{U}$ of the right-hand side of (2.16), namely the first equality in (2.14) is satisfied.

In order to prove the second equality in (2.14) we apply Lemma 2.9, with $Y = \mathcal{U}$, $Z = L^p(\Omega;\mathbb{R}^{n\times n})$, A the gradient operator, $\Phi(t) := -\langle f, t\rangle$, and

$$\Psi(P) := \int_\Omega j((MP)^{\mathrm{sym}})\beta\,dx\,,$$

where, for brevity, we have set $M := D\Psi_\varepsilon^{-T}$ and $\beta := |\det D\Psi_\varepsilon|$. In view of (2.11), we may write

$$-\inf E_\varepsilon = -\inf_{u\in Y}\{\Psi(Au) + \Phi(u)\} = \inf_{\sigma\in Z^*}\{\Psi^*(\sigma) + \Phi^*(-A^*\sigma)\}\,. \tag{2.17}$$

It is easy to see that $-A^*$ acts on Z^* as the divergence operator, while $\Phi^*(\tau)$ equals 0 if $\tau = -f$ in Y^* and $+\infty$ otherwise. Let us compute the Fenchel conjugate of Ψ: given $Q \in L^{p'}(\Omega;\mathbb{R}^{n\times n})$, exploiting the fact that M is a smooth invertible matrix, we have

$$\begin{aligned}
\Psi^*(Q) &= \sup_{P\in L^p(\Omega;\mathbb{R}^{n\times n})}\Big\{\int_\Omega P : Q\,dx - \int_\Omega j((MP)^{\mathrm{sym}})\beta\,dx\Big\}\\
&= \sup_{P\in L^p(\Omega;\mathbb{R}^{n\times n})}\Big\{\int_\Omega MP : M^{-T}Q\,dx - \int_\Omega j((MP)^{\mathrm{sym}})\beta\,dx\Big\}\\
&= \sup_{P\in L^p(\Omega;\mathbb{R}^{n\times n})}\Big\{\int_\Omega P : M^{-T}Q\,dx - \int_\Omega j(P^{\mathrm{sym}})\beta\,dx\Big\}\\
&= \Big(\int_\Omega j(\cdot)\beta\,dx\Big)^*(M^{-T}Q) = \int_\Omega j^*(\beta^{-1}M^{-T}Q)\beta\,dx\,.
\end{aligned}$$

By combining these results with (2.17), we get

$$-\inf E_\varepsilon = \inf_{\sigma\in\mathcal{S}}\Big\{\int_\Omega j^*(\beta^{-1}M^{-T}\sigma)\beta\,dx \;:\; -\operatorname{div}\sigma = f \text{ in } \mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n)\Big\}. \tag{2.18}$$

Recalling that M and β are $D\Psi_\varepsilon^{-T}$ and $|\det D\Psi_\varepsilon|$, respectively, we infer that the right-hand side of (2.18) agrees with the infimum of the functional H_ε defined in (2.13), and the proof is concluded. □

In the following we use $\mathcal{C}^*(f_\varepsilon)$ to denote the equivalent dual formulation of $\mathcal{C}(f_\varepsilon)$, that is, $\mathcal{C}^*(f_\varepsilon) := \inf H_\varepsilon$. This notation is consistent with the one given for $\varepsilon = 0$ (see Lemma 2.1).
We now study the asymptotic behaviour of the sequences $\{E_\varepsilon\}$ and $\{H_\varepsilon\}$ in the limit as $\varepsilon \to 0$. To that aim we need in particular to endow $\mathcal{S}$ with a notion of weak convergence, which is the following one:

$$\sigma_h \rightharpoonup \sigma \iff \begin{cases} \lim_h \sigma_h = \sigma \quad \text{weakly in } L^{p'}(\Omega;\mathbb{R}^{n\times n})\\ \lim_h \operatorname{div}\sigma_h = \operatorname{div}\sigma \quad \text{weakly * in } \mathcal{M}(\overline{\Omega};\mathbb{R}^n)\,. \end{cases} \tag{2.19}$$

Lemma 2.13. *(i) On the space $\mathcal{U}$ endowed with the weak $W^{1,p}$ convergence, the sequence of functionals E_ε in* (2.12) *is equicoercive and, as $\varepsilon \to 0$, it Γ-converges to the functional E_0. In particular, every sequence $u_\varepsilon \in \operatorname{Argmin}(E_\varepsilon)$ admits a subsequence which converges weakly in $W^{1,p}(\Omega;\mathbb{R}^n)$ to some function $u_0 \in \operatorname{Argmin}(E_0)$.*

(ii) On the space $\mathcal{S}$ endowed with weak convergence defined in (2.19)*, the sequence of functionals H_ε in* (2.13) *is equicoercive and, as $\varepsilon \to 0$, it Γ-converges to the functional H_0. In particular,*

every sequence $\sigma_\varepsilon \in \mathrm{Argmin}(H_\varepsilon)$ *admits a subsequence which converges weakly in* $\mathcal{S}$ *to some field* $\sigma_0 \in \mathrm{Argmin}(H_0)$.

Proof. (i) Thanks to the growth assumptions on j made in (2.2), the equicoercivity of the functionals E_ε follows from Korn inequality (see Lemma 2.11 above) and the continuous embedding of $W^{1,p}(\Omega;\mathbb{R}^n)$ into $C(\Omega;\mathbb{R}^n)$. Notice that, in the scalar case when E_ε are defined on $W^{1,p}(\Omega)$, one has just to use, in place of the Korn inequality, either the Poincaré or the Poincaré-Wirtinger inequality, respectively, when either $\Sigma \neq \emptyset$ or $\Sigma = \emptyset$ and f is balanced.
To prove the Γ-liminf inequality, let $u_\varepsilon \rightharpoonup u$ in $W^{1,p}(\Omega;\mathbb{R}^n)$. Possibly passing to a subsequenge, by the compact embedding of $W^{1,p}(\Omega;\mathbb{R}^n)$ into $C(\Omega;\mathbb{R}^n)$, we may assume that $u_\varepsilon \to u$ uniformly on Ω. Thus $\lim_\varepsilon \langle f, u_\varepsilon\rangle = \langle f, u\rangle$.
Moreover, we observe that $D\Psi_\varepsilon^{-1}$ and $\det D\Psi_\varepsilon$ admit the following asymptotic expansions in terms of ε:

$$D\Psi_\varepsilon^{-1} = I - \varepsilon DV + \varepsilon^2 N_\varepsilon\,, \qquad \det D\Psi_\varepsilon = 1 + \varepsilon \operatorname{div} V + \varepsilon^2 m_\varepsilon\,, \tag{2.20}$$

for some N_ε and m_ε which are uniformly bounded in L^∞-norm.
Then the inequality $\liminf_\varepsilon E_\varepsilon(u_\varepsilon) \geq E(u)$ follows from the L^p-weak lower semicontinuity of the functional $z \mapsto \int_\Omega j(z)\,dx$ (*cf.* Lemma 2.10) and the expansions (2.20).
Finally, the Γ-limsup inequality is satisfied by choosing the sequence $u_\varepsilon \equiv u$, thanks to the L^p-strong continuity of the functional $z \mapsto \int_\Omega j(z)\,dx$ (see again Lemma 2.10).
(ii) Thanks to the growth assumptions of order p from both above and below made on j in (2.2), the Fenchel conjugate satisfies analogous conditions with p replaced by its conjugate exponent p'. Then the coercivity and the Γ-convergence in the $L^{p'}$-weak topology are straightforward by arguing similarly as done above for E_ε, and taking into account that the constraint $-\operatorname{div}\sigma = f$ is closed in the weak $L^{p'}$-topology. □

As last ingredient, we recall the classical min-max result:

Lemma 2.14. *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be nonempty convex subsets of two locally convex topological vector spaces, and let* $\mathcal{B}$ *be compact. Assume that* $L : \mathcal{A}\times\mathcal{B} \to \mathbb{R}$ *is such that for every* $b \in \mathcal{B}$, $L(\cdot, b)$ *is convex, and for every* $a \in \mathcal{A}$, $L(a,\cdot)$ *is upper semicontinuous and concave. Then, if the quantity*

$$\gamma := \inf_{a\in\mathcal{A}} \sup_{b\in\mathcal{B}} L(a,b)$$

is finite, we have $\gamma = \sup_{b\in\mathcal{B}} \inf_{a\in\mathcal{A}} L(a,b)$, *and there exists* $b^\star \in \mathcal{B}$ *such that* $\inf_{a\in\mathcal{A}} L(a,b^\star) = \gamma$. *If in addition* $\mathcal{A}$ *is compact and, for every* $b \in \mathcal{B}$, $L(\cdot,b)$ *is lower semicontinuous, there exists* $a^\star \in \mathcal{A}$ *such that* $L(a^\star, b^\star) = \gamma$.

Proof. See [11, p. 263] and [13]. □

We are now in a position to give the

Proof of Theorem 2.2. We prove separately a lower and an upper bound for the differential quotient

$$r_\varepsilon(V) := \frac{\mathcal{C}(f_\varepsilon) - \mathcal{C}(f)}{\varepsilon}\,.$$

Upper bound: Let u_ε and σ be optimal for E_ε and H_0, respectively. In particular, we have $\int_\Omega j^*(\sigma)\,dx < +\infty$, which implies that $\sigma \in L^{p'}(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}})$ (*cf.* (2.3)).
By using Lemmas 2.1 and 2.12, the Fenchel inequality, and the expansions in (2.20), and taking into account that σ is symmetric, we have

$$\begin{aligned}
r_\varepsilon(V) &= \frac{\mathcal{C}(f_\varepsilon) - \mathcal{C}^*(f)}{\varepsilon} = \frac{-E_\varepsilon(u_\varepsilon) - H_0(\sigma)}{\varepsilon} \\
&= \frac{1}{\varepsilon}\Big(\int_\Omega [-j(D\Psi_\varepsilon^{-T} Du_\varepsilon)\,\det D\Psi_\varepsilon - j^*(\sigma)]\,dx + \langle f, u_\varepsilon\rangle \Big) \\
&\leq \frac{1}{\varepsilon}\Big(\int_\Omega [-(D\Psi_\varepsilon^{-T} Du_\varepsilon) : \sigma \det D\Psi_\varepsilon + j^*(\sigma)(\det D\Psi_\varepsilon - 1)]\,dx + \langle f, u_\varepsilon\rangle \Big) \\
&= \frac{1}{\varepsilon}\Big(-\int_\Omega Du_\varepsilon : \sigma\,dx + \langle f, u_\varepsilon\rangle \Big) \\
&\quad + \int_\Omega [Du_\varepsilon : DV\sigma + (j^*(\sigma) - Du_\varepsilon : \sigma)\operatorname{div} V]\,dx + R_\varepsilon\,.
\end{aligned}$$

Here R_ε denotes the following remainder:

$$R_\varepsilon := \varepsilon \int_\Omega \Big[\big(j^*(\sigma) - Du_\varepsilon : \sigma \big) m_\varepsilon + DV^T Du_\varepsilon : \sigma (\operatorname{div} V + \varepsilon m_\varepsilon) - N_\varepsilon^T Du_\varepsilon : \sigma \det D\Psi_\varepsilon \Big]\, dx\,,$$

being m_ε and N_ε as in (2.20). Then the uniform bounds on m_ε and N_ε yield that $R_\varepsilon \to 0$ as $\varepsilon \to 0$. Since $H_0(\sigma)$ is finite, we have $-\operatorname{div} \sigma = f$ in $\mathcal{D}'(\overline{\Omega} \setminus \Sigma; \mathbb{R}^n)$. In particular, this implies

$$-\int_\Omega Du_\varepsilon : \sigma\, dx + \langle f, u_\varepsilon \rangle = 0 \qquad \forall \varepsilon\,.$$

Finally, by Lemma 2.13 (i), up to subsequences $\{u_\varepsilon\}$ converges weakly in $W^{1,p}(\Omega; \mathbb{R}^n)$ to a function $u_0 \in \operatorname{Argmin}(E_0)$, which satisfies the optimality condition (2.5). We conclude that

$$\limsup_\varepsilon r_\varepsilon(V) \le \int_\Omega [Du : DV\sigma - j(Du) \operatorname{div} V]\, dx = \int_\Omega A(u_0, \sigma) : DV\, dx\,,$$

where last equality follows from the symmetry of σ and the definition (2.7) of $A(u)$.

Lower bound: Let σ_ε and u be optimal for H_ε and E_0, respectively. We observe that $D\Psi_\varepsilon$ admits the following asymptotic expansions in terms of ε:

$$D\Psi_\varepsilon = I + \varepsilon DV + \varepsilon^2 M_\varepsilon\,, \tag{2.21}$$

for some M_ε which is uniformly bounded in L^∞-norm.
By using Lemmas 2.1 and 2.12, the Fenchel inequality, the second expansion in (2.20), and (2.21), we obtain

$$\begin{aligned} r_\varepsilon(V) &= \frac{\mathcal{C}^*(f_\varepsilon) - \mathcal{C}(f)}{\varepsilon} = \frac{H_\varepsilon(\sigma_\varepsilon) + E_0(u)}{\varepsilon} \\ &= \frac{1}{\varepsilon} \Big(\int_\Omega [j^*(\det D\Psi_\varepsilon^{-1}\, D\Psi_\varepsilon \sigma_\varepsilon)\, \det D\Psi_\varepsilon + j(e(u))]\, dx - \langle f, u \rangle \Big) \\ &\ge \frac{1}{\varepsilon} \Big(\int_\Omega [Du : D\Psi_\varepsilon \sigma_\varepsilon + j(Du)(1 - \det D\Psi_\varepsilon)]\, dx - \langle f, u \rangle \Big) \\ &= \frac{1}{\varepsilon} \Big(\int_\Omega Du : \sigma_\varepsilon\, dx - \langle f, u \rangle \Big) + \int_\Omega [Du : DV \sigma_\varepsilon - j(Du) \operatorname{div} V]\, dx + \widetilde{R}_\varepsilon\,. \end{aligned}$$

Here $\widetilde{R}_\varepsilon$ denotes the following remainder:

$$\widetilde{R}_\varepsilon = \varepsilon \int_\Omega \Big[Du : M_\varepsilon \sigma_\varepsilon - j(Du) m_\varepsilon \Big]\, dx\,,$$

being m_ε as in (2.20) and M_ε as in (2.21). Similarly as above, it is easily checked that $\widetilde{R}_\varepsilon$ is infinitesimal as $\varepsilon \to 0$. Since $H_\varepsilon(\sigma_\varepsilon)$ is finite, we have $-\operatorname{div} \sigma_\varepsilon = f$ in $\mathcal{M}(\overline{\Omega} \setminus \Sigma; \mathbb{R}^n)$. In particular, this implies

$$-\int_\Omega Du : \sigma_\varepsilon\, dx + \langle f, u \rangle = 0 \qquad \forall \varepsilon\,.$$

Finally, by Lemma 2.13 (ii), up to subsequences $\{\sigma_\varepsilon\}$ converges weakly in $L^{p'}(\Omega; \mathbb{R}^{n\times n})$ to a field $\sigma_0 \in \operatorname{Argmin}(H_0)$. Hence we conclude that

$$\liminf_\varepsilon r_\varepsilon(V) \ge \int_\Omega [Du : DV\sigma - j(Du) \operatorname{div} V]\, dx = \int_\Omega A(u, \sigma_0) : DV\, dx\,.$$

Conclusion: by the arbitrariness of the elements taken in $\operatorname{Argmin}(E_0)$ and $\operatorname{Argmin}(H_0)$, the previous bounds imply that

$$\inf_\sigma \sup_u \int_\Omega A(u, \sigma_0) : DV\, dx \le \liminf_\varepsilon r_\varepsilon(V) \le \limsup_\varepsilon r_\varepsilon(V) \le \sup_u \inf_\sigma \int_\Omega A(u, \sigma_0) : DV\, dx\,, \tag{2.22}$$

where u and σ vary in $\operatorname{Argmin}(E_0)$ and $\operatorname{Argmin}(H_0)$, respectively. Since the sup-inf at the right-hand side of (2.22) is lower than or equal to the inf-sup at the left-hand side, we infer that all the inequalities in (2.22) are in fact equalities, giving (2.6). The existence of a saddle point $(u^\star, \sigma^\star) \in \operatorname{Argmin}(E_0) \times \operatorname{Argmin}(H_0)$ such that $\mathcal{C}'(f, V) = \int_\Omega A(u^\star, \sigma^\star) : DV\, dx$ follows from Lemma 2.14, with $\mathcal{A} = \operatorname{Argmin}(H_0)$, $\mathcal{B} = \operatorname{Argmin}(E_0)$, and $L(a, b) = \int_\Omega A(b, a) : DV\, dx$. The compactness of $\mathcal{A}$ and $\mathcal{B}$ has already been pointed out in the proof of Lemma 2.13. $\square$

3. Derivative of the Wasserstein distance

In this section we prove the existence and a representation formula for the first derivative of the Wasserstein distance with respect to horizontal variations of the load, defined as in (1.5). Throughout the section, we add the assumption that the open bounded set $\Omega \subset \mathbb{R}^n$ is *convex*, and f will be a signed measure with zero average such that f^+ and f^- are probability measures on $\overline{\Omega}$. For every $x, y \in \Omega$, we denote by $[x, y]$ the oriented segment from x to y.
Before stating our result, we collect in the next lemma some well-known facts we need from transport theory (for a proof, we refer for instance to [1] and [7]).

Lemma 3.1. *(i) The Wasserstein distance admits the following dual formulation:*

$$\mathcal{C}_\infty(f) = W_1(f^+, f^-) = \max\{\langle f, u\rangle \ : \ u \in \mathrm{Lip}_1(\Omega)\}\,, \tag{3.1}$$

and a function u which solves (3.1) *is called a Kantorovich potential.*
(ii) Let u be a Kantorovich potential solving (3.1)*, and let γ be an optimal plan for $W_1(f^+, f^-)$, to which we associate the vector measure $\lambda \in \mathcal{M}(\overline{\Omega}; \mathbb{R}^n)$ defined by*

$$\langle \lambda, \Phi\rangle := \int_{\overline{\Omega}\times\overline{\Omega}} \Big[\int_{[x,y]} \Phi \cdot \frac{y-x}{|y-x|}\, d\mathcal{H}^1\Big]\, d\gamma(x,y) \qquad \forall \Phi \in C(\overline{\Omega}; \mathbb{R}^n)\,,$$

whose total variation measure is given by

$$\langle \mu, \varphi\rangle := \int_{\overline{\Omega}\times\overline{\Omega}} \Big[\int_{[x,y]} \varphi\, d\mathcal{H}^1\Big]\, d\gamma(x,y) \qquad \forall \varphi \in C(\overline{\Omega})\,. \tag{3.2}$$

Then $\lambda = (\nabla_\mu u)\mu$ and (u, μ) solves the Monge-Kantorovich equations

$$-\operatorname{div}((\nabla_\mu u)\,\mu) = f\,, \qquad\qquad |\nabla_\mu u| = 1 \quad \mu - a.e.\,, \tag{3.3}$$

where $\nabla_\mu u$ is the tangential gradient of u with respect to μ.
In addition, for γ-a.e. $(x, y) \in \overline{\Omega} \times \overline{\Omega}$, the function u is affine on $[x, y]$ with unit slope, and it holds

$$\nabla_\mu u(z) \equiv \frac{y-x}{|y-x|} \quad \text{for } \mathcal{H}^1\text{-a.e. } z \in [x,y].$$

(iii) If $(u, \mu) \in \mathrm{Lip}_1(\Omega) \times \mathcal{M}^+(\Omega)$ is a solution to (3.3)*, the measure $\lambda = (\nabla_\mu u)\mu$ solves*

$$\inf\Big\{\int |\lambda| \ : \ \lambda \in \mathcal{M}(\overline{\Omega}; \mathbb{R}^n) \ : \ -\operatorname{div}\lambda = f \text{ in } \mathbb{R}^n \setminus \Sigma\Big\}\,. \tag{3.4}$$

Remark 3.2. If we take a function φ which is merely in $L^1_\mu(\overline{\Omega})$, then for γ-a.e. $(x, y) \in \overline{\Omega} \times \overline{\Omega}$ it turns out that φ is defined $\mathcal{H}^1$-a.e. on $[x, y]$, and by an approximation argument the equality (3.2) continues to hold.

We are now ready to state:

Theorem 3.3. *The derivative of $\mathcal{C}_\infty$ at f in direction V defined according to* (1.5) *exists and is given by*

$$\mathcal{C}'_\infty(f, V) = \inf\{L(\gamma, V) \ : \ \gamma \in \operatorname{Argmin} \mathcal{C}_\infty(f)\}\,, \tag{3.5}$$

with

$$L(\gamma, V) := \int_{\overline{\Omega}\times\overline{\Omega}} (V(y) - V(x)) \cdot \frac{y-x}{|y-x|}\, d\gamma(x,y)\,.$$

Remark 3.4. If we disintegrate γ with respect to its marginals $f^\pm$ (namely we write $\gamma = \gamma^+_y \otimes f^+ = \gamma^-_x \otimes f^-$ for suitable families $(\gamma^+_y)_{y\in\overline{\Omega}}$, $(\gamma^-_x)_{x\in\overline{\Omega}}$ of probabilities on $\overline{\Omega}$, see *e.g.* [2]), and we set

$$c^+(x) := \int_{\overline{\Omega}} \frac{y-x}{|y-x|}\, d\gamma^+_x(y)\,, \qquad c^-(y) := \int_{\overline{\Omega}} \frac{y-x}{|y-x|}\, d\gamma^-_y(x)\,,$$

the integral which defines $L(\gamma, V)$ can be recast as

$$L(\gamma, V) = \langle c^- f^- - c^+ f^+, V\rangle = \int_{\overline{\Omega}} V(y) \cdot c^-(y)\, df^-(y) - \int_{\overline{\Omega}} V(x) \cdot c^+(x)\, df^+(x)\,.$$

Proof of Theorem 3.3. Let γ be an optimal plan for $\mathcal{C}_\infty(f)$. For every $\varepsilon > 0$, we denote by γ_ε the measure $(\Psi_\varepsilon \times \Psi_\varepsilon)_\sharp \gamma$ on $\overline{\Omega} \times \overline{\Omega}$. It is easy to check that γ_ε is still a probability on the product space $\overline{\Omega} \times \overline{\Omega}$ and has marginals $(\pi_{1,2})_\sharp \gamma_\varepsilon = (\Psi_\varepsilon)_\sharp f^\pm = f_\varepsilon^\pm$, in particular $\gamma_\varepsilon \in \Pi(f_\varepsilon^+, f_\varepsilon^-)$. Therefore,

$$\limsup_\varepsilon \frac{\mathcal{C}_\infty(f_\varepsilon) - \mathcal{C}_\infty(f)}{\varepsilon} \leq \limsup_\varepsilon \frac{1}{\varepsilon} \int_{\overline{\Omega}\times\overline{\Omega}} \Big(|x + \varepsilon V(x) - y - \varepsilon V(y)| - |x-y| \Big)\, d\gamma(x,y)$$
$$= \int_{\overline{\Omega}\times\overline{\Omega}} (V(y) - V(x)) \cdot \frac{y-x}{|y-x|}\, d\gamma(x,y) = L(\gamma, V)\,. \tag{3.6}$$

Here in order to pass to the limit as $\varepsilon \to 0^+$, we have applied the dominated convergence theorem, since the family

$$h_\varepsilon(x,y) := \frac{|x + \varepsilon V(x) - y - \varepsilon V(y)| - |x-y|}{\varepsilon} \tag{3.7}$$

is uniformly bounded by the constant $2\max_{\overline{\Omega}} |V|$ and, as $\varepsilon \to 0$, it converges pointwise to

$$h(x,y) := (V(y) - V(x)) \cdot \frac{y-x}{|y-x|} \quad \forall (x,y) \in \overline{\Omega} \times \overline{\Omega}\,,$$

(intended to be zero on the diagonal $x = y$). Notice that, since h_ε are equi-Lipschitz, the pointwise convergence yields in fact the uniform convergence on compact sets. By the arbitrariness of γ in $\operatorname{Argmin} \mathcal{C}_\infty(f)$ in (3.6), we conclude that

$$\limsup_\varepsilon \frac{\mathcal{C}_\infty(f_\varepsilon) - \mathcal{C}_\infty(f)}{\varepsilon} \leq \inf\{L(\gamma, V) \ :\ \gamma \in \operatorname{Argmin} \mathcal{C}_\infty(f)\}\,. \tag{3.8}$$

As already pointed out at the beginning of the proof, the sets $\Pi(f^+, f^-)$ and $\Pi(f_\varepsilon^+, f_\varepsilon^-)$ have a one-to-one correspondence, so that we may rewrite $\mathcal{C}_\infty(f_\varepsilon)$ and $\mathcal{C}(f)$ as

$$\mathcal{C}_\infty(f_\varepsilon) = \inf_{\gamma \in \mathcal{P}(\overline{\Omega}\times\overline{\Omega})} \mathcal{F}_\varepsilon(\gamma)\,, \qquad \mathcal{C}_\infty(f) = \inf_{\gamma \in \mathcal{P}(\overline{\Omega}\times\overline{\Omega})} \mathcal{F}(\gamma)\,,$$

with

$$\mathcal{F}_\varepsilon(\gamma) = \int_{\overline{\Omega}\times\overline{\Omega}} |x + \varepsilon V(x) - y - \varepsilon V(y)|\, d\gamma(x,y) + \chi(\gamma)\,,$$
$$\mathcal{F}(\gamma) = \int_{\overline{\Omega}\times\overline{\Omega}} |x-y|\, d\gamma(x,y) + \chi(\gamma)\,,$$

being χ the indicatrix function of $\Pi(f^+, f^-)$ (which equals 0 if $\gamma \in \Pi(f^+, f^-)$ and $+\infty$ otherwise). It is easy to check that $\mathcal{F}_\varepsilon$ Γ-converges to $\mathcal{F}$ in the metric space of probability measures $\mathcal{P}(\overline{\Omega} \times \overline{\Omega})$ endowed with the weak topology. In particular, a minimizing sequence γ_ε of $\mathcal{C}_\infty(f_\varepsilon)$ converges, up to subsequences, to an optimal plan $\overline{\gamma}$ of $\mathcal{C}_\infty$. In view of the optimality of γ_ε, we may write

$$\liminf_\varepsilon \frac{\mathcal{C}_\infty(f_\varepsilon) - \mathcal{C}_\infty(f)}{\varepsilon} \geq \liminf_\varepsilon \int_{\overline{\Omega}\times\overline{\Omega}} h_\varepsilon(x,y)\, d\gamma_\varepsilon(x,y)\,, \tag{3.9}$$

where the integrand h_ε is defined in (3.7). Eventually, since $h_\varepsilon \to h$ uniformly on compact sets, and $\gamma_\varepsilon \rightharpoonup \overline{\gamma}$, we obtain

$$\liminf_\varepsilon \int_{\overline{\Omega}\times\overline{\Omega}} h_\varepsilon(x,y)\, d\gamma_\varepsilon(x,y) = \liminf_\varepsilon \int_{\overline{\Omega}\times\overline{\Omega}} h(x,y)\, d\gamma_\varepsilon(x,y) \geq \int_{\overline{\Omega}\times\overline{\Omega}} h(x,y)\, d\overline{\gamma}(x,y)\,. \tag{3.10}$$

By combining (3.9) and (3.10), recalling that $\overline{\gamma}$ is an optimal plan for $\mathcal{C}_\infty(f)$, we infer that

$$\liminf_\varepsilon \frac{\mathcal{C}_\infty(f_\varepsilon) - \mathcal{C}_\infty(f)}{\varepsilon} \geq \inf\{L(\gamma, V) \ :\ \gamma \in \operatorname{Argmin} \mathcal{C}_\infty(f)\}\,.$$

This inequality, together with (3.8), gives (3.5). □

Proposition 3.5. *For any $\gamma \in \operatorname{Argmin} \mathcal{C}_\infty(f)$, if u is a Kantorovich potential and $\mu = \mu(\gamma)$ is the measure defined in* (3.2)*, we can rewrite $L(\gamma, V)$ as*

$$L(\gamma, V) = \int_{\overline{\Omega}} (\nabla_\mu u \otimes \nabla_\mu u) : DV\, d\mu\,. \tag{3.11}$$

In particular, if $\operatorname{Argmin} \mathcal{C}_\infty(f)$ *contains a unique transport plan* $\overline{\gamma}$*, we have the following distributional equality*

$$\mathcal{C}'_\infty(f,V) = -\langle \operatorname{div}((\nabla_\mu u \otimes \nabla_\mu u)\mu), V\rangle \qquad \forall V \in \mathcal{D}(\Omega;\mathbb{R}^n)\,, \tag{3.12}$$

with u *a Kantorovich potential and* $\mu = \mu(\overline{\gamma})$*.*

Proof. By taking $\varphi = (\nabla_\mu u \otimes \nabla_\mu u) : DV$ in (3.2) (*cf.* Remark 3.2), we get

$$\int_{\overline{\Omega}} (\nabla_\mu u \otimes \nabla_\mu u) : DV\, d\mu = \int_{\overline{\Omega}\times\overline{\Omega}} \Big[\int_{[x,y]} (\tau_{xy} \otimes \tau_{xy}) : DV\, d\mathcal{H}^1 \Big]\, d\gamma(x,y)\,. \tag{3.13}$$

Exploiting the parametrization $[0,1] \ni t \mapsto (1-t)x + ty$ of the segment $[x,y]$, whose velocity has modulus $|y-x|$, we obtain

$$\begin{aligned}
\int_{[x,y]} (\tau_{xy} \otimes \tau_{xy}) : DV\, d\mathcal{H}^1 &= \int_0^1 \Big[\Big(\frac{(y-x)}{|y-x|} \otimes \frac{(y-x)}{|y-x|} \Big) : DV((1-t)x+ty) \Big]\, |y-x|\, dt \\
&= \int_0^1 \frac{(y-x)}{|y-x|} \cdot \Big(DV((1-t)x+ty)(y-x) \Big)\, dt = \int_0^1 \frac{(y-x)}{|y-x|} \cdot \frac{d}{dt} \Big(V((1-t)x+ty) \Big)\, dt \\
&= \frac{(y-x)}{|y-x|} \cdot \Big(V(y) - V(x) \Big)\,.
\end{aligned} \tag{3.14}$$

By combining (3.13) and (3.14), we conclude that

$$\int_{\overline{\Omega}} (\nabla_\mu u \otimes \nabla_\mu u) : DV\, d\mu = \int_{\overline{\Omega}\times\overline{\Omega}} (V(y) - V(x)) \cdot \frac{y-x}{|y-x|}\, d\gamma(x,y)\,,$$

namely the assertion (3.11).
Finally, in case there exists a unique transport plan, (3.12) readily follows by combining (3.5) and (3.11). □

Example 3.6. Let $f^+ = \delta_A$ and $f^- = \delta_B$. Then the unique transport plan is $\overline{\gamma} = \delta_A \otimes \delta_B$, and we have

$$\mathcal{C}'_\infty(f,V) = \int_{\overline{\Omega}\times\overline{\Omega}} (V(y) - V(x)) \cdot \frac{y-x}{|y-x|}\, d(\delta_A \otimes \delta_B) = \int_{\overline{\Omega}} V(y) \cdot c^-(y)\, d\delta_B(y) - \int_{\overline{\Omega}} V(x) \cdot c^+(x)\, d\delta_A(x)\,,$$

with

$$c^-(y) = \frac{y-A}{|y-A|}\,, \qquad c^+(x) = \frac{B-x}{|B-x|}\,.$$

Example 3.7. Let Ω be a convex planar domain with perimeter 1, let $P \in \Omega$, and let $f = \mathcal{H}^1 \llcorner \partial\Omega - \delta_P$. The unique transport plan is $\overline{\gamma} := \delta_P \otimes \mathcal{H}^1 \lfloor_{\partial\Omega}$; moreover, by taking $d_P(y) := |y-P|$ as a test function in (3.1) of $\mathcal{C}_\infty(f)$, we see that $\mathcal{C}_\infty(f) = \int_{\partial\Omega} |y-P|\, d\mathcal{H}^1(y)$, so that d_P is a Kantorovich potential. Hence

$$\begin{aligned}
\mathcal{C}'_\infty(f,V) &= \Big(\int_{\partial\Omega} \frac{P-x}{|P-x|}\, d\mathcal{H}^1(x) \Big) V(P) - \int_{\partial\Omega} V(x) \cdot \frac{P-x}{|P-x|}\, d\mathcal{H}^1(x) \\
&= \int_{\overline{\Omega}} V(y) \cdot c^-(y)\, d\delta_P(y) - \int_{\overline{\Omega}} V(x) \cdot c^+(x)\, d(\mathcal{H}^1 \llcorner \partial\Omega)(x)\,,
\end{aligned}$$

with

$$c^-(y) = \int_{\partial\Omega} \frac{y-x}{|y-x|}\, \mathcal{H}^1(x)\,, \qquad c^+(x) = \frac{P-x}{|P-x|}\,.$$

In particular, we have

$$\mathcal{C}'_\infty(f,V) = \langle \Big(\int_{\partial\Omega} \frac{P-x}{|P-x|}\, d\mathcal{H}^1(x) \Big)\, \delta_P, V\rangle \qquad \forall V \in \mathcal{D}(\Omega;\mathbb{R}^2)\,.$$

In view of Proposition 3.5, we deduce that

$$\Big(\int_{\partial\Omega} \frac{P-x}{|P-x|}\, d\mathcal{H}^1(x) \Big)\, \delta_P = -\operatorname{div}((\nabla_\mu u \otimes \nabla_\mu u)\mu) \qquad \text{in } \mathcal{D}'(\Omega;\mathbb{R}^2)\,. \tag{3.15}$$

Actually, the above equality can be checked by direct computations, by using the explicit expressions of μ and $\nabla_\mu u$.

Denoting by ρ and θ the polar coordinates centered at P, we have

$$\mu = \frac{\alpha(\theta)}{\rho}\,\mathcal{L}^2 \llcorner \Omega\,, \qquad \nabla_\mu u = e_\theta := (\cos\theta, \sin\theta) \tag{3.16}$$

where

$$\alpha(\theta) := \sqrt{(\overline{\rho}(\theta))^2 + (\overline{\rho}'(\theta))^2}\,, \qquad \overline{\rho}(\theta) := \max\{\rho \in \mathbb{R}^+ \,:\, \rho e_\theta \in \overline{\Omega}\}\,.$$

(To obtain (3.16) one has to apply the equality (3.2) in the proof of Theorem 3.3).
Using the expressions in (3.16), some easy computations show that $\operatorname{div}((\nabla_\mu u \otimes \nabla_\mu u)\mu) = 0$ in $\mathbb{R}^2 \setminus \{P\}$. On the other hand, let $\varphi \in \mathcal{D}(K;\mathbb{R}^2)$, where K is a compact neighbourhood of P. Setting for brevity $M := (e_\theta \otimes e_\theta)\frac{\alpha(\theta)}{\rho}$, and denoting by ν the unit outer normal to $K \setminus B_r(P)$, we get

$$\begin{aligned}
&-\langle \operatorname{div}((\nabla_\mu u \otimes \nabla_\mu u)\mu), \varphi\rangle = \lim_{r\to 0} \int_{K\setminus B_r(P)} M : D\varphi\, dx \\
&= \lim_{r\to 0} \Big[-\int_{K\setminus B_r(P)} \operatorname{div} M \cdot \varphi\, dx + \int_{\partial K} \langle M\nu, \varphi\rangle\, d\mathcal{H}^1 + \int_{\partial B_r(P)} \langle M\nu, \varphi\rangle\, d\mathcal{H}^1 \Big] \\
&= \lim_{r\to 0} \int_{\partial B_r(P)} \langle M\nu, \varphi\rangle\, d\mathcal{H}^1 = -\lim_{r\to 0} \int_0^{2\pi} \alpha(\theta) e_\theta \cdot \varphi(r e_\theta)\, d\theta \\
&= -\varphi(P) \cdot \int_0^{2\pi} \alpha(\theta) e_\theta\, d\theta = \langle \Big(\int_{\partial\Omega} \frac{P-x}{|P-x|}\, d\mathcal{H}^1(x) \Big)\, \delta_P, \varphi \rangle\,,
\end{aligned}$$

concluding the proof of (3.15).

Example 3.8. When the transport plan is not unique, the derivative $\mathcal{C}'_\infty(f,V)$ may be not linear with respect to V. For instance, let Ω be the square $(0,1)\times(0,1)$, let $A=(0,0)$, $B=(1,0)$, $C=(1,1)$, and $D=(0,1)$ be the corners of the square, and take $f^+ := (\delta_A+\delta_C)/2$ and $f^- := (\delta_B+\delta_D)/2$. It is easy to see that $\mathcal{C}_\infty(f)=1$, and that the optimal transport plans are of the form

$$\gamma_s(x,y) := \frac{1}{2}\delta_A(x) \otimes (s\delta_B(y) + (1-s)\delta_D(y)) + \frac{1}{2}\delta_C(x) \otimes ((1-s)\delta_B(y) + s\delta_D(y))\,, \tag{3.17}$$

with $s \in [0,1]$. Rearranging the terms above, we infer that every optimal transport plan is a convex combination of γ_0 and γ_1, that is $\gamma_s = s\gamma_0 + (1-s)\gamma_1$. In view of formula (3.5), we have

$$\begin{aligned}
\mathcal{C}'_\infty(f,V) &= \inf_{s\in[0,1]} L(\gamma_s, V) = \inf_{s\in[0,1]} L(s\gamma_0 + (1-s)\gamma_1, V) = \min\{L(\gamma_0,V), L(\gamma_1,V)\} \\
&= \frac{1}{2}\min\Big\{ \big(V(D)-V(A)\big)\cdot \frac{D-A}{|D-A|} + \big(V(B)-V(C)\big)\cdot\frac{B-C}{|B-C|}\,, \\
&\qquad \big(V(B)-V(A)\big)\cdot\frac{B-A}{|B-A|} + \big(V(D)-V(C)\big)\cdot\frac{D-C}{|D-C|} \Big\}
\end{aligned} \quad ,$$

which is clearly not linear with respect to V.

4. Approximation result

We now turn to the approximation formula (1.6). As in the previous section, we still assume that Ω is a convex domain, and that f a signed measure with zero average such that f^+ and f^- are probability measures on $\overline{\Omega}$.
From now on, we use the notation $\mathcal{C}_p(f)$ for the scalar compliance defined according to (2.8), where we assuming without further mention that the Dirichlet region Σ is empty and the space of admissible displacements is $W^{1,p}(\Omega)$.

Theorem 4.1. *If $\mathcal{C}_\infty(f)$ admits a unique optimal transport plan $\overline{\gamma}$, for every admissible V we have*

$$\lim_{p\to+\infty} \mathcal{C}'_p(f,V) = \mathcal{C}'_\infty(f,V) = L(\overline{\gamma}, V)\,.$$

Remark 4.2. Without the uniqueness assumption, the equality $\lim_{p\to+\infty}\mathcal{C}'_p(f,V)=\mathcal{C}'_\infty(f,V)$ can be false. Indeed, in view of formula (2.9), the map $V\mapsto\mathcal{C}'_p(f,V)$ is linear. Thus, the validity of the previous equality entails the linearity of the map $V\mapsto\mathcal{C}'_\infty(f,V)$. When there exist multiple transport plans, such linearity be violated, as shown by Example 3.8. It is an open problem to establish whether or not the limit of $\mathcal{C}'_p(f,V)$ as $p\to+\infty$ exists also in case of multiple transport plans; a positive answer would imply the existence of a peculiar transport plan $\widehat{\gamma}$ such that $\lim_{p\to+\infty}\mathcal{C}'_p(f,V)=L(\widehat{\gamma},V)$. This is the case for instance in Example 3.8, where by symmetry it is readily seen that $\widehat{\gamma}=\gamma_{1/2}$ (the latter plan being defined according to (3.17)). If such a plan $\widehat{\gamma}$ do exist in general, then a further interesting question would be to discover some selection principle (of "entropy type"), allowing to detect it among the family of all optimal transport plans.

In order to prove Theorem 4.1, we need to recall some notions of convergence with respect to measures (see [6, §3] and references therein).

Definition 4.3. Let $\mu_k\,,\mu\in\mathcal{M}^+(\overline{\Omega})$, with $\mu_k\rightharpoonup\mu$, and let $g_k\,,\ g:\overline{\Omega}\to\mathbb{R}^d$ be μ_k and μ-integrable, respectively. We say that g_k weakly converges to g with respect to the pair (μ_k,μ), and we write $g_k\overset{w(\mu_k,\mu)}{\longrightarrow}g$, if $g_k\mu_k\rightharpoonup g\mu$ in $\mathcal{M}(\overline{\Omega};\mathbb{R}^d)$.
We say that g_k strongly converges to g with respect to the pair (μ_k,μ), and we write $g_k\overset{s(\mu_k,\mu)}{\longrightarrow}g$, if for every $\varphi\in C_0(\mathbb{R}^n\times\mathbb{R}^d)$ we have

$$\lim_{k\to+\infty}\int_{\overline{\Omega}}\varphi(x,g_k(x))\,d\mu_k=\int_{\overline{\Omega}}\varphi(x,g(x))\,d\mu\,.$$

If $h_k\,,\ h:\overline{\Omega}\to\mathbb{R}^m$ are μ_k and μ-integrable, respectively, we say that the pair (g_k,h_k) strongly-weakly converges to (g,h) with respect to (μ_k,μ), and we write $(g_k,h_k)\overset{sw(\mu_k,\mu)}{\longrightarrow}(g,h)$, if $g_k\overset{s(\mu_k,\mu)}{\longrightarrow}g$ and $h_k\overset{w(\mu_k,\mu)}{\longrightarrow}h$.

Lemma 4.4. *If $(g_k,h_k)\in L^\alpha_{\mu_k}(\overline{\Omega};\mathbb{R}^n)\times L^\beta_{\mu_k}(\overline{\Omega};\mathbb{R}^n)$ strongly-weakly converge to (g,h) with respect to (μ_k,μ), then the product sequence $\{g_k\cdot h_k\}$ converges to $g\cdot h$ weakly with respect to (μ_k,μ), provided that $\frac{1}{\alpha}+\frac{1}{\beta}<1$, and $\|g_k\|_{L^\alpha_{\mu_k}(\overline{\Omega})}\le C$, $\|h_k\|_{L^\beta_{\mu_k}(\overline{\Omega})}\le C$ for some constant $C>0$.*

Proof. See [6, Remark 3.8]. □

Lemma 4.5. *As $p\to+\infty$, up to (not relabeled) subsequences, we have*

$$\begin{aligned}
&u_p\to u \textit{ uniformly };\\
&|\nabla u_p|^p \textit{ is uniformly bounded in } L^1(\Omega)\,; &(4.1)\\
&\mu_p:=|\nabla u_p|^{p-2}\mathcal{L}^n\rightharpoonup\mu \textit{ in } \mathcal{M}^+(\overline{\Omega})\,; &(4.2)\\
&\mu_p\nabla u_p\rightharpoonup\mu\nabla_\mu u \textit{ in } \mathcal{M}(\overline{\Omega};\mathbb{R}^n)\,; &(4.3)\\
&\nabla u_p\overset{s(\mu_p,\mu)}{\longrightarrow}\nabla_\mu u\,, &(4.4)
\end{aligned}$$

for some (u,μ) solution to (3.3).

Proof. See [4, Theorem 4.1 and Remark 4.1]. □

Now we establish a preliminary result needed for the proof of Theorem 4.1.

Lemma 4.6. *In the limit as $p\to+\infty$, up to a (not relabeled) subsequence, we have*

$$\frac{\nabla u_p}{|\nabla u_p|}\overset{w(\mu_p,\mu)}{\longrightarrow}\nabla_\mu u\,. \tag{4.5}$$

Moreover, for every $\Psi\in C(\mathbb{R}^n;\mathbb{R}^{n\times n})$, we have

$$\lim_{p\to+\infty}\int_\Omega\langle\Psi\nabla u_p,\nabla u_p\rangle\,d\mu_p=\lim_{p\to+\infty}\int_\Omega\Big\langle\Psi\nabla u_p,\frac{\nabla u_p}{|\nabla u_p|}\Big\rangle\,d\mu_p\,. \tag{4.6}$$

Proof. For brevity, let us denote by α_p the modulus $|\nabla u_p|$. In view of the weak convergence of ∇u_p with respect to (μ_p,μ) (see (4.3)), in order to prove (4.5) it is enough to show that

$(\nabla u_p - \nabla u_p/\alpha_p)$ weakly converges to zero with respect to (μ_p, μ): for every test function $\varphi \in C_0(\mathbb{R}^n;\mathbb{R}^n)$, we have

$$\begin{aligned}\Big|\int_\Omega \varphi\cdot(\nabla u_p - \nabla u_p/\alpha_p)\,d\mu_p\Big| &= \Big|\int_\Omega \varphi\cdot\frac{\nabla u_p}{\alpha_p}(\alpha_p-1)\alpha_p^{p-2}\,dx\Big| \\ &\le C\int_\Omega |\alpha_p^{p-1}-\alpha_p^{p-2}|\,dx\,. \end{aligned}\tag{4.7}$$

In order to conclude, we need to remove the modulus in the integral. To this aim, we exploit the inequality $t^{q'}/q' \ge t - 1/q$, for every $t>0$ and $1/q+1/q'=1$ (*cf.* pag. 20 in [4], just after formula (35)): taking $t=\alpha_p^{p-2}$ and $q=p-1$, we get

$$\begin{aligned}&q' = \frac{p-1}{p-2}\,,\quad \frac{q'}{q}=\frac{1}{p-2}\,,\\ &t^{q'} = \alpha_p^{p-1}\,,\\ &\alpha_p^{p-1} - q'\alpha_p^{p-2} + \frac{q'}{q} \ge 0\,,\end{aligned}\tag{4.8}$$

so that, by the triangle inequality,

$$|\alpha_p^{p-1}-\alpha_p^{p-2}| \le \alpha_p^{p-1}-\alpha_p^{p-2} + 2\frac{q'}{q}\,.$$

Inserting this inequality in (4.7), we get

$$\int_\Omega |\alpha_p^{p-1}-\alpha_p^{p-2}|\,dx \le \int_\Omega \alpha_p^{p-1}\,dx - \int_\Omega \alpha_p^{p-2}\,dx + \frac{2|\Omega|}{p-2}\,.\tag{4.9}$$

Again by (4.8), exploiting the fact that $q'\to 1$ and $q'/q\to 0$ as $p\to+\infty$, we infer that

$$\limsup_p \int_\Omega \alpha_p^{p-2}\,dx \le \limsup_p \int_\Omega \alpha_p^{p-1}\,dx = \mu(\overline{\Omega})\,,\tag{4.10}$$

where the last equality can be found in [4, Lemma 3.2 and proof of Theorem 4.1]. On the other hand, by the weak convergence $\mu_p \rightharpoonup \mu$ and the lower semicontinuity of the mass, we have

$$\mu(\overline{\Omega}) \le \liminf_p \mu_p(\Omega) = \liminf_p \int_\Omega \alpha_p^{p-2}\,dx\,.\tag{4.11}$$

By combining (4.10) and (4.11), we infer that

$$\lim_p \int_\Omega \alpha_p^{p-2} = \lim_p \int_\Omega \alpha_p^{p-1} = \mu(\overline{\Omega})\,.\tag{4.12}$$

Eventually, by combining (4.7) with (4.9) and (4.12), we conclude that

$$\Big|\int_\Omega \varphi\cdot\Big(\nabla u_p - \frac{\nabla u_p}{\alpha_p}\Big)\,d\mu_p\Big| \to 0 \quad \text{as } p\to+\infty\,,$$

which, by the arbitrariness of φ, gives (4.5). With an analogous argument it is easy to prove (4.6). □

Proof of Theorem 4.1. In view of (2.9) we have

$$\mathcal{C}_p'(f,V) = \int_\Omega A(u_p) : DV\,dx\,.$$

By the uniform L^1 boundedness of $|\nabla u_p|^p$ (see (4.1)), up to passing to a (not relabeled) subsequence, we may assume that the limit as $p\to+\infty$ of $\mathcal{C}_p'(f,V)$ exists.
We claim that, up to subsequences,

$$\mathcal{C}_p'(f,V) \to \int_{\overline{\Omega}} (\nabla_\mu u \otimes \nabla_\mu u) : DV\,d\mu\,,\tag{4.13}$$

for some (u,μ) solution to (3.3).
Recalling the expression (2.10) of $A(u_p)$ and the definition (4.2) of μ_p, we may write

$$A(u_p) = |\nabla u_p|^{p-2}(\nabla u_p\otimes\nabla u_p) - \frac{|\nabla u_p|^p}{p}I\,.$$

Again by (4.1) and by (4.6) in Lemma 4.6 with $\Psi = DV$, we get

$$\lim_p \mathcal{C}_p'(f,V) = \lim_p \int_\Omega \langle DV\nabla u_p, \nabla u_p\rangle \, d\mu_p = \lim_p \int_\Omega \Big\langle DV\nabla u_p, \frac{\nabla u_p}{|\nabla u_p|}\Big\rangle \, d\mu_p \,.$$

According to Definition 4.3, by (4.4) in Lemma 4.5 and (4.5) in Lemma 4.6, we infer that $(DV\nabla u_p, \nabla u_p/|\nabla u_p|)$ strongly-weakly converges to $(DV\nabla_\mu u, \nabla_\mu u)$ with respect to (μ_p,μ). Moreover, $DV\nabla u_p$ and $\nabla u_p/|\nabla u_p|$ are uniformly bounded in $L^2_{\mu_p}(\Omega;\mathbb{R}^n)$ and $L^q_{\mu_p}(\Omega;\mathbb{R}^n)$, for every $q \in (2,+\infty)$, respectively. Therefore, by applying Lemma 4.4 with $g_k = DV\nabla u_p$, $h_k = \nabla u_p/|\nabla u_p|$, $g = DV\nabla_\mu u$, $h = \nabla_\mu u$, $\alpha = 2$, $\beta = q$, we get

$$\lim_p \int_\Omega \Big\langle DV\nabla u_p, \frac{\nabla u_p}{|\nabla u_p|}\Big\rangle \, d\mu_p = \int_\Omega \langle DV\nabla_\mu u, \nabla_\mu u\rangle \, d\mu \,,$$

concluding the proof of (4.13).

We observe that our uniqueness assumption for the optimal transport plan $\overline{\gamma}$, combined with Theorem 4.13 in [19], implies that problem (3.4) admits a unique solution $\overline{\lambda}$. Then, by Lemma 3.1 (iii), we can rewrite the r.h.s. of (4.13) as

$$\int_\Omega (\nabla_\mu u \otimes \nabla_\mu u) : DV \, d\mu = \int_\Omega \Big(\frac{d\overline{\lambda}}{d|\overline{\lambda}|} \otimes \frac{d\overline{\lambda}}{d|\overline{\lambda}|}\Big) : DV d|\overline{\lambda}| = L(\overline{\gamma}, V) \,.$$

Since this limit is independent of the choice of the subsequence, we have proved that the whole sequence $\mathcal{C}_p'(f,V)$ converges to $L(\overline{\gamma},V)$, which by Theorem 3.3 agrees with $\mathcal{C}_\infty'(f,V)$. □

5. Appendix

Proof of (1.2). We proceed in the case where Σ has positive $\mathcal{H}^{N-1}$ measure, and we set for brevity $j(z) = \frac{1}{p}|z|^p$. Let $f_\varepsilon, f \in W^{-1,p'}(\overline{\Omega}\setminus\Sigma)$ be such that $g_\varepsilon = \frac{f_\varepsilon - f}{\varepsilon}$ converges strongly to g. Let u_p be the optimal displacement for $\mathcal{C}_p(f)$. By definition of $\mathcal{C}_p(f_\varepsilon)$, it holds

$$\mathcal{C}_p(f_\varepsilon) \geq \langle f_\varepsilon, u_p\rangle - \int_\Omega j(\nabla u_p)\, dx \,,$$

and hence, by definition of u_p, we have

$$\liminf_{\varepsilon\to 0} \frac{\mathcal{C}_p(f_\varepsilon) - \mathcal{C}_p(f)}{\varepsilon} \geq \; \liminf_{\varepsilon\to 0} \langle g_\varepsilon, u_p\rangle \; = \; \langle g, u_p\rangle \,.$$

To obtain the converse inequality, we exploit the dual formulation (2.4) of the compliance.
Let $\sigma_p := \nabla j(\nabla u_p)$, which is optimal for $\mathcal{C}_p^*(f)$, and let τ be any vector field in $\mathcal{S}$ such that $-\operatorname{div}\tau = g$ in $\mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n)$. By the strong convergence $g_\varepsilon \to g$ in $W^{-1,p'}(\overline{\Omega}\setminus\Sigma)$, we can pick up a sequence τ_ε such that $\tau_\varepsilon \to \tau$ in $L^{p'}(\Omega;\mathbb{R}^n)$ and $-\operatorname{div}\tau_\varepsilon = g_\varepsilon$ in $\mathcal{D}'(\overline{\Omega}\setminus\Sigma;\mathbb{R}^n)$. Then by definition of $\mathcal{C}_p^*(f_\varepsilon)$ and of σ_p, it holds

$$\mathcal{C}_p^*(f_\varepsilon) - \mathcal{C}_p^*(f) \leq \int_\Omega j^*(\sigma_p + \varepsilon\tau_\varepsilon) - j^*(\sigma_p)\, dx \leq \int_\Omega j^*(\sigma_p + \varepsilon\tau) - j^*(\sigma_p)\, dx + C\varepsilon\|\tau_\varepsilon - \tau\|_{L^{p'}(\Omega;\mathbb{R}^n)} \,,$$

for some positive constant C independent of ε. For the last inequality, we have used the estimate

$$\big||z_1|^{p'} - |z_2|^{p'}\big| \leq C(1 + |z_1|^{p'-1} + |z_2|^{p'-1})|z_1 - z_2| \,,$$

valid in $\mathbb{R}^n$ (see, *e.g.*, [15, Proposition 4.64]). Moreover, $\sigma \mapsto \int j^*(\sigma)$ is continuous in $L^{p'}(\Omega;\mathbb{R}^n)$ and its subdifferential at σ is the singleton $\nabla j^*(\sigma)$; in particular, the integral functional is Gateaux-differentiable (*cf.* [12]). Hence, recalling that $\tau_\varepsilon \to \tau$ in $L^{p'}(\Omega;\mathbb{R}^n)$ and $\nabla j^*(\sigma_p) = \nabla u_p$, we get

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \frac{\mathcal{C}_p^*(f+\varepsilon g) - \mathcal{C}_p^*(f)}{\varepsilon} \quad &\leq \limsup_{\varepsilon\to 0} \int_\Omega \frac{j^*(\sigma_p+\varepsilon\tau) - j^*(\sigma_p)}{\varepsilon}\, dx + \limsup_{\varepsilon\to 0} \|\tau_\varepsilon - \tau\|_{L^{p'}(\Omega;\mathbb{R}^n)} \\
&= \int_\Omega \nabla j^*(\sigma_p)\cdot\tau\, dx = \int_\Omega \nabla u_p \cdot \tau\, dx = -\langle \operatorname{div}\tau, u_p\rangle = \langle g, u_p\rangle \,.
\end{aligned}$$

□

References

[1] L. Ambrosio: Lecture notes on optimal transport problems, *Mathematical aspects of evolving interfaces*, 1–52, Lecture Notes in Math., 1812, Springer Verlag, Berlin, 2003.

[2] L. Ambrosio, N. Fusco, D. Pallara: Functions of Bounded Variation and Free Discontinuity Problems, *Oxford Mathematical Monographs*, Clarendon Press, Oxford, 2000.

[3] G. Bouchitté: Convex analysis and duality methods. Variational Thechniques, *Encyclopedia of Mathematical physics*, 642–652, Academic Press, 2006.

[4] G. Bouchitté, G. Buttazzo, L. De Pascale: A p-Laplacian approximation for some mass optimization problems, *J. Optimiz. Theory App.*, **118**, no. 1, 1–25 (2003).

[5] G. Bouchitté, G. Buttazzo, L. De Pascale: The Monge-Kantorovich problem for distributions and applications, *J. Convex Anal.*, **17**, no. 3-4, 925–943 (2010).

[6] G. Bouchitté, G. Buttazzo, I. Fragalà : Convergence of Sobolev spaces on varying manifolds, *J. Geom. Anal.*, **11**, no. 3, 399–422 (2001).

[7] G. Bouchitté, G. Buttazzo, P. Seppecher : Shape optimization solutions via Monge-Kantorovich equation, *Comptes Rendus de l'Academie de Sciences de Paris*, **324**, no. 10, 1185–1191 (1997).

[8] G. Bouchitté, T. Champion, C. Jimenez: Completion of the space of measures in the Kantorovich norm, *Riv. Mat. Univ. Parma*, **4**, 127–139 (2005).

[9] G. Bouchitté, I. Fragalà, I. Lucardesi : Shape derivatives for minima of integral functionals, *Math. Programm.*, **148**, 111–142 (2014).

[10] G. Bouchitté, I. Fragalà, I. Lucardesi : A variational method for second order shape derivatives, *SIAM J. Control Optim.*, **54**, 1056-1084 (2016).

[11] F.H. Clarke: Multiple integrals of Lipschitz functions in the calculus of variations, *Proc. Amer. Math. Soc.*, 260–264 (1977).

[12] I. Ekeland: Convexity Methods in Hamiltonian Mechanics, Springer-Verlag, 1990.

[13] I. Ekeland: Théorie des Jeux, Univ. France, Paris, 1975.

[14] M. Delfour, J.P. Zolésio: Shapes and geometries. Analysis, differential calculus, and optimization. Advances in Design and Control, SIAM, Philadelpia, 2001.

[15] I. Fonseca, G. Leoni: Modern Methods in Calculus of Variations: L^p spaces. Springer Monographs in Mathematics, Springer, 2007.

[16] Tröltzsch, Fredi: Optimal control of partial differential equations. AMS Monographs in Mathematics , Graduate Studies in Mathematics, Volume 112, 2010.

[17] A. Henrot, M. Pierre: Variation et Optimisation de Formes. Une Analyse Géométrique. Mathématiques & Applications, **48**, Springer, Berlin, 2005.

[18] N. Kikuchi, J.T. Oden: Contact problems in elasticity: a study of variational inequalities and finite elements methods, SIAM Studies in Applied Mathematics, **8**, SIAM Philadelphia, 1988.

[19] F. Santambrogio: Optimal transport for applied mathematicians. Calculus of variations, PDEs, and modeling. Progress in Nonlinear Differential Equations and their Applications, **87**, Birkhäuser/Springer, Cham, 2015.

[20] J. Sokolowski, J.P. Zolésio: Differential stability of solutions to unilateral problems. Free boundary problems: application and theory, Vol. IV (Maubuisson, 1984), 537-547, *Res. Notes in Math.*, **121**, Pitman, Boston, 1985.

♯ Laboratoire IMATH, Université de Toulon, 83957 La Garde Cedex (France)

∗ Dipartimento di Matematica, Politecnico di Milano, Piazza L. da Vinci, 20133 Milano (Italy)

♭ Institut Élie Cartan de Lorraine, BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex (France)