RESEARCH

Inner and Outer Versions of Hyper-Elasticity

Pablo Pedregal[1]

Abstract

Through suitable changes of variables for a typical problem in hyper-elasticity, either in the reference or deformed configurations, one can setup and analyze versions of the same problem in terms of inner or outer maps or variations. Though such kind of transformations are part of the classical background in the Calculus of Variations, we explore under what sets of hypotheses such versions can be shown to have minimizers and be equivalent to the standard form of the problem. Such sets of hypotheses lead naturally to some distinct poly-convex energy densities for hyper-elasticity. Likewise we explore optimality in either of the two forms through a special way to generate one-parameter families of feasible deformations, feasibility including injectivity and non-interpenetration of matter.

1 Introduction

We would like to look at the central problem in hyper-elasticity, that of minimizing an internal energy functional of the integral form

$$E(\boldsymbol{u}) = \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x})) \, d\boldsymbol{x},$$

from the viewpoint of inner- and/or outer-deformations. Here Ω is a Lipschitz, bounded domain in $\mathbb{R}^N$, the density

$$W(\mathbf{F}) : \mathbb{R}_+^{N \times N} \to \mathbb{R}, \quad \det \mathbf{F} > 0,$$

is a continuous density taking the value $+\infty$ when $\det \mathbf{F} \leq 0$, and mappings

$$\boldsymbol{u}(\boldsymbol{x}) : \Omega \to \mathbb{R}^N$$

belong to a suitable class of feasible Sobolev deformations in $H^1(\Omega; \mathbb{R}^N)$. The energy density W should comply with some further conditions that we will overlook for the time being. In addition, inhomogeneity, i.e. some explicit dependence $W(\boldsymbol{x}, \mathbf{F})$ on $\boldsymbol{x} \in \Omega$, might be

---

✉ P. Pedregal

1   Departamento de Matemáticas, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain

treated as well. Additional terms accounting for bulk and surface forces can be added in a standard form. The boundary $\partial\Omega$ is a Lipschitz $(N-1)$-hyper-surface that is divided, as usual, in three parts: Γ_0, where a displacement boundary condition is imposed; Γ_t, where the usual natural-boundary, applied traction-free condition is assumed; and a possible additional negligible (with respect to the $N-1$-Hausdorff measure in $\partial\Omega$) set $\mathbb{N}$. In this way, we have

$$\partial\Omega = \Gamma_0 \cup \Gamma_t \cup \mathbb{N}. \tag{1.1}$$

For definiteness, we will take $N = 3$, though most of our results can be appropriately extended for other cases as well. In addition, to facilitate our discussion, we will stick to a global condition of place in which

$$\Gamma \equiv \Gamma_0 \equiv \partial\Omega, \quad \Gamma_t = \mathbb{N} = \emptyset.$$

For future reference, we adopt the following definition.

Definition 1.1 A continuous density

$$W(\mathbf{F}) : \mathbb{R}^{3\times3} \to \mathbb{R} \cup \{+\infty\},$$

such that

$$\{W < +\infty\} = \mathbb{R}_+^{3\times3} \equiv \{\mathbf{F} \in \mathbb{R}^{3\times3} : \det\mathbf{F} > 0\}, \tag{1.2}$$

and

$$W(\mathbf{R}\,\mathbf{F}) = W(\mathbf{F}), \quad \mathbf{F} \in \mathbb{R}_+^{3\times3}, \mathbf{R} \in \mathcal{SO}(3), \tag{1.3}$$

is called a density for hyper-elasticity.

Condition (1.3) is the usual frame indifference to ensure the mechanical meaning of our internal energy density W. We have not placed other demands on such densities to stress the fact that from the point of view of Analysis, the main trouble is related to how such densities take on the value $+\infty$. Note that such an integrand cannot admit a global polynomial upper bound. This is part of the principal difficulty of hyper-elasticity.

Suppose $\boldsymbol{u}_\circ$ is a bi-Lipschitz mapping between two Lipschitz, bounded domains

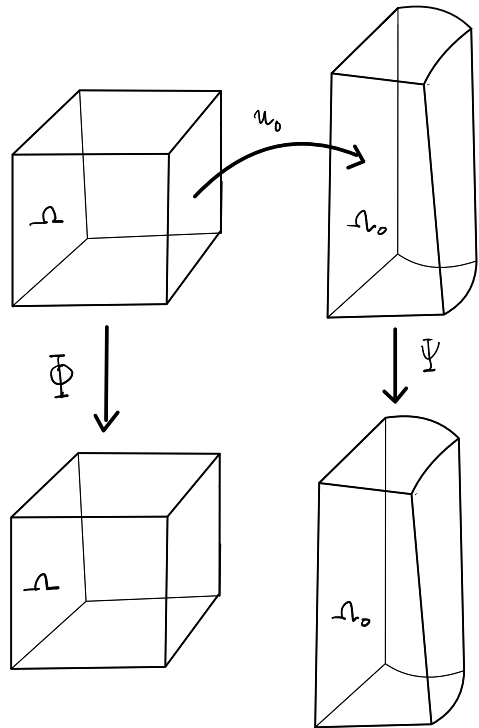$$\Omega, \quad \Omega_\circ = \boldsymbol{u}_\circ(\Omega),$$

in such a way that

$$\boldsymbol{u}_\circ \in W^{1,\infty}(\Omega; \Omega_\circ), \quad \boldsymbol{u}_\circ^{-1} \in W^{1,\infty}(\Omega_\circ; \Omega). \tag{1.4}$$

This mapping $\boldsymbol{u}_\circ$ will be used to setup boundary conditions on Γ. Though this is an important restriction, it can, on the other hand, cover many situations of interest, so it enjoys a certain degree of generality. The map $\boldsymbol{u}_\circ$ with these properties is fixed once and for all.

We would like to look at the central problem in hyper-elasticity in the form

$$I(\Psi, \Phi) = \int_\Omega W(\nabla U_{\Psi,\Phi}(\boldsymbol{x}))\,d\boldsymbol{x}, \quad U_{\Psi,\Phi}(\boldsymbol{x}) = \Psi(\boldsymbol{u}_\circ(\Phi^{-1}(\boldsymbol{x}))),$$

Fig. 1 Mappings Ψ and Φ

where mappings $\Phi(\boldsymbol{x}) \in H^1(\Omega; \mathbb{R}^3)$ are such that (see Fig. 1)

$$\Phi(\boldsymbol{x}) = \boldsymbol{x} \text{ on } \Gamma, \tag{1.5}$$

while $\Psi(\boldsymbol{y}) \in H^1(\Omega_\circ; \mathbb{R}^3)$ should comply with

$$\Psi(\boldsymbol{y}) = \boldsymbol{y} \text{ on } \Gamma_\circ \equiv \boldsymbol{u}_\circ(\Gamma) = \partial\Omega_\circ. \tag{1.6}$$

In fact, the simultaneously use of Φ and Ψ is redundant, given that $\boldsymbol{u}_\circ$ and its inverse $\boldsymbol{u}_\circ^{-1}$ are well-defined, Lipschitz maps between Ω and $\Omega_\circ$.

Remark 1.1 Conditions (1.5) and/or (1.6) can be relaxed to deal eventually with pure traction or mixed problems. See Remark 4.4 below.

We will consider the three intimately related variational problems

$$\text{Minimize in } \boldsymbol{u} \in \mathcal{A}_+ : \quad E(\boldsymbol{u}) = \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x})) \, d\boldsymbol{x}, \tag{1.7}$$

$$\text{Minimize in } \Phi \in \mathcal{I}_+ : \quad I(\Phi) = \int_\Omega W(\nabla(\boldsymbol{u}_\circ \circ \Phi^{-1})(\boldsymbol{x})) \, d\boldsymbol{x}, \tag{1.8}$$

$$\text{Minimize in } \Psi \in \mathcal{O}_+ : \quad O(\Psi) = \int_\Omega W(\nabla(\Psi \circ \boldsymbol{u}_\circ)(\boldsymbol{x})) \, d\boldsymbol{x}, \tag{1.9}$$

where the respective feasible classes of mappings are

$$\mathcal{A}_+ = \{\boldsymbol{u} \in H^1(\Omega; \mathbb{R}^3) : \boldsymbol{u} = \boldsymbol{u}_\circ \text{ on } \Gamma, \boldsymbol{u} \text{ is one-to-one a.e. in } \Omega, \tag{1.10}$$

$$\det \nabla \boldsymbol{u} > 0 \text{ a.e. in } \Omega\}$$

$$\mathcal{I}_+ = \{\Phi \in H^1(\Omega; \mathbb{R}^3) : \Phi = \mathbf{id} \text{ on } \Gamma, \Phi \text{ is one-to-one a.e. in } \Omega, \tag{1.11}$$

$$\det \nabla \Phi > 0 \text{ a.e. in } \Omega\}$$

$$\mathcal{O}_+ = \{\Psi \in H^1(\Omega_\circ; \mathbb{R}^3) : \Psi = \mathbf{id} \text{ on } \Gamma_\circ, \Psi \text{ is one-to-one a.e. in } \Omega_\circ, \tag{1.12}$$

$$\det \nabla \Psi > 0 \text{ a.e. in } \Omega_\circ\}.$$

It is easy to realize that problems (1.8)-(1.11) and (1.9)-(1.12) are subproblems of (1.7)-(1.10) because, at least formally, maps of the form $\boldsymbol{u}_\circ \circ \Phi^{-1}$ and $\Psi \circ \boldsymbol{u}_\circ$ belong to $\mathcal{A}_+$; but suspect that under suitable assumptions all three ought to be equivalent. Our goal is to see to what extent we can learn something from looking at the initial problem in hyperelasticity in either of these two forms. Specifically, we seek to:

(1) establish hypotheses under which problems (1.8)-(1.11) and (1.9)-(1.12) admit minimizers;
(2) refine those sets of conditions to ensure that all three problems are equivalent with the natural rules to pass among minimizers;
(3) explore optimality conditions for the inner and outer versions based on one-parameter families of deformations belonging entirely to either $\mathcal{I}_+$ or $\mathcal{O}_+$.

The Dirichlet boundary condition $\Phi(\boldsymbol{x}) = \boldsymbol{x}$ on $\Gamma \equiv \partial\Omega$ preserves the global condition of place that is to be respected through the given mapping $\boldsymbol{u}_\circ$. Since the class of maps Φ that we would like to consider are to be one-to-one and onto, they are legitimate changes of variables, and then it is elementary to rewrite the internal energy functional to be minimized in the form

$$\int_\Omega W(\nabla \boldsymbol{u}_0(\boldsymbol{x}) \nabla \Phi(\boldsymbol{x})^{-1}) \det \nabla \Phi(\boldsymbol{x}) \, d\boldsymbol{x}.$$

We thus have two intimately-connected energy functionals

$$E(\boldsymbol{u}) = \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x})) \, d\boldsymbol{x}, \tag{1.13}$$

$$I(\Phi) = \int_\Omega W(\nabla \boldsymbol{u}_0(\boldsymbol{x}) \nabla \Phi(\boldsymbol{x})^{-1}) \det \nabla \Phi(\boldsymbol{x}) \, d\boldsymbol{x}, \tag{1.14}$$

where competing maps for the first are

$$\boldsymbol{u}(\boldsymbol{x}) : \Omega \to \mathbb{R}^3, \quad \boldsymbol{u} = \boldsymbol{u}_\circ \text{ on } \partial\Omega,$$

whereas for the second they are

$$\Phi(\boldsymbol{x}) : \Omega \to \mathbb{R}^3, \quad \Phi = \mathbf{id} \text{ on } \partial\Omega.$$

Roughly, as indicated above, the former class of maps $\boldsymbol{u}$ is larger than that of mappings of the form $\boldsymbol{u}_\circ \circ \Phi^{-1}$ for Φ belonging to the latter. Existence results for the first are well-known. Our intention is then three-fold:

- we would like to prove some standard existence results for the second (1.14)-(1.11);
- examine hypotheses under which the two problems are equivalent in the sense that the rule

$$\boldsymbol{u} \mapsto \boldsymbol{u}_{\circ} \circ \Phi^{-1}$$

can be used to go from minimizers for the first to minimizers of the second, and viceversa;
- explore necessary conditions of optimality.

The natural idea is to enforce the usual hypotheses on the integrand

$$\overline{W}(\boldsymbol{x}, \mathbf{X}) = W(\nabla \boldsymbol{u}_0(\boldsymbol{x}) \mathbf{X}^{-1}) \det \mathbf{X} \tag{1.15}$$

to guarantee the existence of minimizers in appropriate functional spaces, and see how these conditions are translated into the original integrand W. Optimality condition for the corresponding variational problem with integrand $\overline{W}$ will also be explored.

As a matter of fact, the transformation (after replacing $\nabla \boldsymbol{u}_{\circ}(\boldsymbol{x})$ by the identity to avoid inhomogeneity),

$$W(\mathbf{X}) \mapsto \mathbb{T}W(\mathbf{X}) \equiv \det \mathbf{X}\, W(\mathbf{X}^{-1}), \tag{1.16}$$

is an involution in the class of densities for hyper-elasticity. It is indeed the rule that is leading our analysis in the following sense:

(1) Poly-convexity turns out to be stable under this transformation.
(2) Coercivity is formulated through suitable lower bounds that are also stable under the same transformation. Indeed, every time that we have a suitable lower bound of the form

$$W(\mathbf{X}) \geq C w(\mathbf{X}), \quad w \geq 0, C > 0,$$

the function

$$C(w(\mathbf{X}) + \mathbb{T}w(\mathbf{X})) \tag{1.17}$$

becomes a valid lower bound both for W and for $\mathbb{T}W$, possibly for a different positive constant $C > 0$, if more matrices are involved as in (1.15).

Our goal is to treat all of these informal ideas in a rigorous way focusing on the main issues associated with any family of variational problems:

(1) existence:
   (a) polyconvexity;
   (b) coercivity;
(2) equivalence of standard, inner and outer versions;
(3) optimality.

We will handle these points in detail for the inner version (1.8)-(1.11), and since the outer version is simpler if anything, we will afterwards indicate the changes for (1.9)-(1.12). Note that the use of inner and outer variations is a classic technique in the Calculus of Variations (see for instance the treatise [11], or [13], [14]). In the context of hyper-elasticity, they were indicated, at the level of optimality, in [5] without the injectivity and the interpenetration-of-matter conditions. See also [4].

2 Existence

We focus on the inner version of the original problem in the form

$$\text{Minimize in } \Phi \in \mathcal{I}_+ : \quad I(\Phi) = \int_\Omega W(\nabla(\boldsymbol{u}_\circ \circ \Phi^{-1})(\boldsymbol{x})) \, d\boldsymbol{x}, \tag{2.1}$$

where $\mathcal{I}_+$ is defined by

$$\mathcal{I}_+ = \{\Phi \in H^1(\Omega; \mathbb{R}^3) : \Phi = \mathbf{id} \text{ on } \Gamma, \, \Phi \text{ is one-to-one a.e. in } \Omega. \tag{2.2}$$

$$\det \nabla \Phi > 0 \text{ a.e. in } \Omega\}.$$

As discussed earlier, this is the inner-version of the standard problem in hyper-elasticity that consists in

$$\text{Minimize in } \boldsymbol{u} \in \mathcal{A}_+ : \quad E(\boldsymbol{u}) = \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x})) \, d\boldsymbol{x}, \tag{2.3}$$

over the class

$$\mathcal{A}_+ = \{\boldsymbol{u} \in H^1(\Omega; \mathbb{R}^3) : \boldsymbol{u} = \boldsymbol{u}_\circ \text{ on } \Gamma, \, \boldsymbol{u} \text{ is one-to-one a.e. in } \Omega, \tag{2.4}$$

$$\det \nabla \boldsymbol{u} > 0 \text{ a.e. in } \Omega\}.$$

We assume that the integrand W is a valid density for hyper-elasticity in the sense of Definition 1.1. As briefly indicated earlier the change of variables $\boldsymbol{x} = \Phi(\boldsymbol{y})$, formally transforms the functional $I(\Phi)$ in (2.1) into

$$I(\Phi) = \int_\Omega W(\nabla \boldsymbol{u}_\circ(\boldsymbol{x}) \nabla \Phi(\boldsymbol{x})^{-1}) \det \nabla \Phi(\boldsymbol{x}) \, d\boldsymbol{x},$$

and hence, we would like to study the variational problem

$$\text{Minimize in } \Phi \in \mathcal{I}_+ : \quad I(\Phi) = \int_\Omega W(\nabla \boldsymbol{u}_\circ(\boldsymbol{x}) \nabla \Phi(\boldsymbol{x})^{-1}) \det \nabla \Phi(\boldsymbol{x}) \, d\boldsymbol{x} \tag{2.5}$$

where

$$\mathcal{I}_+ = \{\Phi \in H^1(\Omega; \mathbb{R}^3) : \Phi = \mathbf{id} \text{ on } \Gamma, \det \nabla \Phi > 0 \text{ a.e. in } \Omega\}. \tag{2.6}$$

Note, due to results cited below about invertibility, that feasible maps in $\mathcal{I}_+$ are one-to-one.

It is by now a well-established fact that there are two fundamental ingredients to showing the existence of global minimizers in the context of hyper-elasticity: poly-convexity, and appropriate coercivity. The statement below (Theorem 2.1) is a classical global existence result in hyper-elasticity ([7]) that we have chosen for definiteness, and because this is the form in which transformation (1.16) becomes more transparent and symmetric. It is a particular case of a more general theorem as a result of fundamental facts mainly in [2] and [15]. See also [10].

Concerning invertibility of Sobolev functions, recent results, particularly in [12] and [15], permit to extend the existence of minimizers from the class

$$\mathcal{A} = \{\boldsymbol{u} \in W^{1,1}(\Omega; \mathbb{R}^3) : E(\boldsymbol{u}) < +\infty, \boldsymbol{u} = \boldsymbol{u}_\circ \text{ on } \Gamma\}$$

to the fundamental subclass for hyper-elasticity

$$\mathcal{A}_+ = \{u \in H^1(\Omega; \mathbb{R}^3) : u = u_\circ \text{ on } \Gamma, u \text{ is one-to-one a.e. in } \Omega, \det \nabla u > 0 \text{ a.e. in } \Omega\}.$$

These results generalize earlier classic theorems in [3, 7, 19]. As a matter of fact, these are sufficient in our setting, once the regularity properties assumed on the mapping $u_\circ$ in (1.4) are considered, to prove that the two classes $\mathcal{A}$ and $\mathcal{A}_+$ are the same set of mappings

$$\mathcal{A} = \mathcal{A}_+ = \{u \in H^1(\Omega; \mathbb{R}^3) : u = u_\circ \text{ on } \Gamma, \det \nabla u > 0 \text{ a.e. in } \Omega\}. \tag{2.7}$$

Note that $\mathcal{A}_+$ is non-empty because $u_\circ \in \mathcal{A}_+$, and $E(u_\circ) < +\infty$.

Theorem 2.1 *Suppose the integrand*

$$W(x, \mathbf{X}) : \Omega \times \mathbb{R}^{3 \times 3} \to \mathbb{R}$$

is a Carathéodory integrand for which:

(1) *for a.e. $x \in \Omega$, the function $W(x, \cdot)$ is a polyconvex density for hyperelasticity*;
(2) *there is a positive constant C, and a exponent $r > 1$, such that for a.e. $x \in \Omega$,*

$$W(x, \mathbf{X}) \geq C(|\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r);$$

and the mapping $u_\circ$ is as indicated before (1.4). Then there is a global minimizer u in the class $\mathcal{A}_+$ in (2.7) for the functional

$$E(u) = \int_\Omega W(x, \nabla u(x)) \, dx.$$

This theorem ensures that there are global minimizers $u \in \mathcal{A}_+$ for problem (2.3).

The point is that we would like to apply this general result to the special integrand

$$\overline{W}(x, \mathbf{X}) = \det \mathbf{X} \, W(\nabla u_\circ(x)\mathbf{X}^{-1}),$$

coming from functional (2.5). In particular, we want to understand how the poly-convexity and the coercivity with respect to $\mathbf{X}$ of such integrand $\overline{W}(x, \mathbf{X})$ translate into suitable properties for W itself.

2.1 Polyconvexity

The poly-convexity condition motivates the following definition. At the level of quasiconvexity, it was introduced and considered in [18].

Definition 2.1 A density for hyperelasticity is declared as inner-poly-convex if for every fixed matrix $\mathbf{F} \in \mathbb{R}_+^{3 \times 3}$, the function

$$W_{\mathbf{F}}(\mathbf{X}) = \det \mathbf{X} \, W(\mathbf{F}\mathbf{X}^{-1}) \tag{2.8}$$

is poly-convex in $\mathbf{X}$.

It turns out that the class of inner-poly-convex functions under (1.2) is the same as the class of poly-convex integrands. Said differently, the class of poly-convex integrands under (1.2), is stable under the transformation (2.8).

Proposition 2.2 *The classes of poly-convex integrands under* (1.2) *and inner-poly-convex integrands under* (1.2) *are exactly the same.*

The key to the proof of this proposition is the following elementary lemma.

Lemma 2.3 *The function*

$$w(\mathbf{A}, \mathbf{B}, t) : \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R}_+ \to \mathbb{R}, \quad \mathbb{R}_+ = \{t \in \mathbb{R} : t > 0\},$$

is convex if and only if the family of functions

$$w_{\mathbf{F}, \mathbf{G}, s}(\mathbf{A}, \mathbf{B}, t) = t\, w\left(\frac{1}{t}\mathbf{F}\mathbf{B}, \frac{1}{t}\mathbf{A}\mathbf{G}, \frac{s}{t}\right) \tag{2.9}$$

are convex for every fixed triplet

$$(\mathbf{F}, \mathbf{G}, s) \in \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R}_+.$$

Proof Since the map

$$(\mathbf{A}, \mathbf{B}) \in \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R}_+ \mapsto (\mathbf{F}\mathbf{B}, \mathbf{A}\mathbf{G}, s)$$

is linear for fixed matrices $\mathbf{F}, \mathbf{G}$, and positive number s, it suffices to check that the operation

$$(\mathbf{x}, t) \in \mathbb{R}^N \times \mathbb{R}_+ \mapsto t\, \omega\left(\frac{1}{t}(\mathbf{x}, 1)\right), \quad w : \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R},$$

remains convex if the function ω is. This is elementary since the combination

$$(\alpha t_1 + (1-\alpha)t_0)\, \omega\left(\frac{1}{\alpha t_1 + (1-\alpha)t_0}(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_0, 1)\right)$$

for $\alpha \in [0, 1]$, $t_0, t_1 > 0$, and $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{R}^N$, can be rewritten as

$$(\alpha t_1 + (1-\alpha)t_0)\, \omega\left(\frac{\alpha t_1}{\alpha t_1 + (1-\alpha)t_0}(\frac{1}{t_1}\mathbf{x}_1, \frac{1}{t_1}) + \frac{(1-\alpha)t_0}{\alpha t_1 + (1-\alpha)t_0}(\frac{1}{t_0}\mathbf{x}_0, \frac{1}{t_0})\right),$$

and the convexity of ω leads directly to the larger quantity (because $(\alpha t_1 + (1-\alpha)t_0) > 0$),

$$\alpha t_1 \omega\left(\frac{1}{t_1}(\mathbf{x}_1, 1)\right) + (1-\alpha)t_0 \omega\left(\frac{1}{t_0}(\mathbf{x}_0, 1)\right).$$

This is the claimed convexity of the family of functions in (2.9).

Conversely, suppose that all functions in (2.9) are convex. In particular, for the choice

$$\mathbf{F} = \mathbf{G} = \mathbf{id}, \quad s = 1, \tag{2.10}$$

the function

$$\gamma(\mathbf{B}, \mathbf{A}, t) = t\, w\left(\frac{1}{t}\mathbf{A}, \frac{1}{t}\mathbf{B}, \frac{1}{t}\right)$$

is convex. By what we have just shown above, the function

$$(\mathbf{A}, \mathbf{B}, t) \mapsto t \, \gamma \left(\frac{1}{t}\mathbf{B}, \frac{1}{t}\mathbf{A}, \frac{1}{t} \right) = w(\mathbf{A}, \mathbf{B}, t)$$

must be convex. □

Proof of Proposition 2.2 A density $W(\mathbf{X})$ is polyconvex if there is a convex function

$$w(\mathbf{A}, \mathbf{B}, t) : \mathbb{R}^{3 \times 3} \times \mathbb{R}^{3 \times 3} \times \mathbb{R}_+ \to \mathbb{R}$$

with

$$W(\mathbf{X}) = w(\mathbf{X}, \operatorname{adj}\mathbf{X}, \det\mathbf{X}).$$

If we now examine the functions $W_\mathbf{F}(\mathbf{X})$ given in (2.8), we see that

$$W_\mathbf{F}(\mathbf{X}) = \det\mathbf{X} W(\mathbf{F}\mathbf{X}^{-1}) = \det\mathbf{X} \, w\left(\mathbf{F}\mathbf{X}^{-1}, \operatorname{adj}(\mathbf{F}\mathbf{X}^{-1}), \det(\mathbf{F}\mathbf{X}^{-1})\right).$$

This can also be recast in the form, using elementary formulae in Linear Algebra,

$$W_\mathbf{F}(\mathbf{X}) = \det\mathbf{X} \, w\left(\frac{1}{\det\mathbf{X}}\mathbf{F}\operatorname{adj}\mathbf{X}, \frac{1}{\det\mathbf{X}}\mathbf{X}\operatorname{adj}\mathbf{F}, \frac{1}{\det\mathbf{X}}\det\mathbf{F}\right).$$

Formally

$$W_\mathbf{F}(\mathbf{X}) = w_{\mathbf{F}, \operatorname{adj}\mathbf{F}, \det\mathbf{F}}(\mathbf{X}, \operatorname{adj}\mathbf{X}, \det\mathbf{X}).$$

By Lemma 2.3, the convexity of $w(\mathbf{A}, \mathbf{B}, t)$ is equivalent to the convexity of the functions

$$w_{\mathbf{F}, \operatorname{adj}\mathbf{F}, \det\mathbf{F}}(\mathbf{A}, \mathbf{B}, t).$$

Notice that for $\mathbf{F} = \mathbf{id}$, we have $\operatorname{adj}\mathbf{F} = \mathbf{id}$ and $\det\mathbf{F} = 1$ as required in (2.10). □

2.2 Coercivity

The other fundamental ingredient is coercivity that according to Theorem 2.1 can be assumed in the form

$$W(\mathbf{F}) \geq C(|\mathbf{F}|^2 + |\operatorname{adj}\mathbf{F}|^2 + \det\mathbf{F}^r), \quad C > 0, r > 1, \mathbf{F} \in \mathbb{R}^{3 \times 3}. \tag{2.11}$$

Note how this lower bound is compatible with (1.2), but it does not enforce the infinity value when $\det\mathbf{F} \leq 0$ because this lower bound is polynomial. Similar, appropriate lower bounds can be assumed in general dimension N for suitable exponents.

A few scratch computations show that the suitable coercivity for the integrand in functional I in (2.5) ought to be

$$W(\mathbf{X}) \geq C\left(\frac{1}{\det\mathbf{X}}|\mathbf{X}|^2 + \frac{1}{\det\mathbf{X}}|\operatorname{adj}\mathbf{X}|^2 + \frac{1}{\det\mathbf{X}^{r-1}}\right), \tag{2.12}$$

for some positive constant $C > 0$.

Lemma 2.4 *Let $\boldsymbol{u}_\circ(\boldsymbol{x}) : \Omega \to \boldsymbol{u}_\circ(\Omega)$ be a Lipschitz homeomorphism. If $W(\mathbf{X})$ is a valid density for hyper-elasticity for which* (2.12) *holds, then there is a positive constant $\tilde{C}$ (depending on $\boldsymbol{u}_\circ$ and C in* (2.12)*) and exponent $r > 1$ such that*

$$\det \mathbf{X}\, W(\nabla \boldsymbol{u}_\circ(\boldsymbol{x})\mathbf{X}^{-1}) \geq \tilde{C}(|\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r),$$

for a.e. $\boldsymbol{x} \in \Omega$.

Proof The proof is elementary. Put

$$\mathbb{H} = \{\nabla \boldsymbol{u}_\circ(\boldsymbol{x}) : \boldsymbol{x} \in \Omega\},$$

a compact set in $\mathbb{R}_+^{3\times3}$. Suppose $\mathbf{F} \in \mathbb{H}$ is fixed. Then it is clear from (2.12) that

$$W(\mathbf{Y}) \geq C \left(\frac{1}{\det \mathbf{Y}} |\mathbf{Y}|^2 + \frac{1}{\det \mathbf{Y}} |\operatorname{adj} \mathbf{Y}|^2 + \frac{1}{\det \mathbf{Y}^{r-1}} \right), \quad \mathbf{Y} = \mathbf{F}\mathbf{X}^{-1},$$

and we immediately find that

$$W(\mathbf{Y})\frac{\det \mathbf{F}}{\det \mathbf{Y}} \geq C \det \mathbf{F} \left(\left| \frac{1}{\det \mathbf{Y}} \mathbf{Y} \right|^2 + \left| \frac{1}{\det \mathbf{Y}} \operatorname{adj} \mathbf{Y}^T \right|^2 + \frac{1}{\det \mathbf{Y}^r} \right)$$

$$= C \det \mathbf{F} \left(|\operatorname{adj}(\mathbf{Y}^{-1})|^2 + |\mathbf{Y}^{-1}|^2 + \det(\mathbf{Y}^{-1})^r \right).$$

On the other hand, from the identities

$$\operatorname{adj}(\mathbf{Y}^{-1}) = (\operatorname{adj} \mathbf{F})^{-1} \operatorname{adj} \mathbf{X}, \quad \mathbf{Y}^{-1} = \mathbf{X}\mathbf{F}^{-1},$$

and the property

$$|\mathbf{A}\mathbf{B}^1| \geq |\mathbf{A}|\,|\mathbf{B}|^{-1},$$

valid for square matrices $\mathbf{A}$, $\mathbf{B}$, it is elementary to arrive at

$$W(\mathbf{Y})\frac{\det \mathbf{F}}{\det \mathbf{Y}} \geq C \det \mathbf{F} \left(|\operatorname{adj} \mathbf{F}|^{-2}|\operatorname{adj} \mathbf{X}|^2 + |\mathbf{F}|^{-2}|\mathbf{X}|^2 + \det \mathbf{F}^{-r} \det \mathbf{X}^r \right).$$

If $\mathbf{F}$ is allowed to move on compact set $\mathbb{H}$, we can definitely find a new positive constant $\tilde{C} > 0$ (depending on C and $\mathbb{H}$) such that

$$W(\mathbf{Y})\frac{\det \mathbf{F}}{\det \mathbf{Y}} \geq \tilde{C} \left(|\operatorname{adj} \mathbf{X}|^2 + |\mathbf{X}|^2 + \det \mathbf{X}^r \right),$$

that is to say

$$\det \mathbf{X}\, W(\nabla \boldsymbol{u}_0(\boldsymbol{x})\mathbf{X}^{-1}) \geq \tilde{C}(|\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r),$$

for a.e. $\boldsymbol{x} \in \Omega$, as desired. $\qquad\square$

2.3 Main Result

As a straightforward outcome of our analysis of poly-convexity and coercivity, the proof of the following theorem is straightforward after Theorem 2.1.

Theorem 2.5 *Let $\boldsymbol{u}_\circ(\boldsymbol{x}) : \Omega \to \boldsymbol{u}_\circ(\Omega)$ be a Lipschitz homeomorphism. If $W(\mathbf{X})$ is a valid density for hyper-elasticity for which we know that*

(1) *W is polyconvex*;
(2) *there is a positive constant $C > 0$ and exponent $r > 1$ such that*

$$W(\mathbf{X}) \geq C \left(\frac{1}{\det \mathbf{X}} |\mathbf{X}|^2 + \frac{1}{\det \mathbf{X}} |\operatorname{adj} \mathbf{X}|^2 + \frac{1}{\det \mathbf{X}^{r-1}} \right), \quad \mathbf{X} \in \mathbb{R}_+^{3 \times 3};$$

then variational problem (2.5)-(2.6) *admits minimizers.*

3 Equivalence of Both Versions

We have stated above basic existence theorems for minimizers of functional E in (2.3) over the class $\mathcal{A}_+$ in (2.7), and for functional I in (2.5) over the set $\mathcal{I}_+$ in (2.6), under suitable sets of assumptions. One would expect that, under hypotheses enabling so, one could go from minimizers of the first to minimizers of the second, and viceversa under the formal rules

$$\boldsymbol{u} = \boldsymbol{u}_\circ \circ \Phi^{-1}, \quad \Phi = \boldsymbol{u}^{-1} \circ \boldsymbol{u}_\circ, \tag{3.1}$$

We explore one such natural set of assumptions in this section.

Since the passage from a mapping $\boldsymbol{u} \in \mathcal{A}_+$ to a mapping $\Phi \in \mathcal{I}_+$, and viceversa involves the inverses of $\boldsymbol{u}$ and Φ through rule (3.1), we need to face some standard conditions guaranteeing that such inverses are Sobolev mappings. This is not a difficult issue under suitable hypotheses. They are reminiscent of conditions in [3].

Lemma 3.1 *Let $\boldsymbol{v} \in W^{1,p}(\Omega; \mathbb{R}^N)$ be a one-to-one mapping a.e. in Ω with $\det \nabla \boldsymbol{v} > 0$ a.e. in Ω, where $\Omega \subset \mathbb{R}^N$ is a Lipschitz, bounded domain, $1 \leq p < \infty$, and $\boldsymbol{v}(\Omega) \subset \mathbb{R}^N$ is another Lipschitz, bounded domain. Suppose*

$$\boldsymbol{v} : \partial \Omega \to \boldsymbol{v}(\partial \Omega) = \partial \boldsymbol{v}(\Omega)$$

is also one-to-one. If

$$\int_\Omega |\nabla \boldsymbol{v}(\boldsymbol{x})^{-1}|^p \det \nabla \boldsymbol{v}(\boldsymbol{x}) \, d\boldsymbol{x} < \infty,$$

then $\boldsymbol{v}^{-1} \in W^{1,p}(\boldsymbol{v}(\Omega); \mathbb{R}^N)$.

Proof Since

$$\boldsymbol{v}^{-1}(\boldsymbol{y}) : \boldsymbol{v}(\Omega) \to \Omega, \quad \boldsymbol{v}^{-1} : \partial \boldsymbol{v}(\Omega) \to \partial \Omega,$$

are well-defined, it suffices to check that its distributional derivatives belong to $L^p(\boldsymbol{v}(\Omega))$. Indeed, by the same change of variables that we have been using several times so far $\boldsymbol{y} = \boldsymbol{v}(\boldsymbol{x})$,

$$\int_{\boldsymbol{v}(\Omega)} |\nabla \boldsymbol{v}^{-1}(\boldsymbol{y})|^p \, d\boldsymbol{y} = \int_{\Omega} |\nabla \boldsymbol{v}^{-1}(\boldsymbol{v}(\boldsymbol{x}))|^p \det \nabla \boldsymbol{v}(\boldsymbol{x}) \, d\boldsymbol{x}.$$

This is precisely the integral in the statement, because

$$\nabla \boldsymbol{v}^{-1}(\boldsymbol{v}(\boldsymbol{x})) = \nabla \boldsymbol{v}(\boldsymbol{x})^{-1}. \qquad \square$$

The invertibility of Sobolev functions is a quite delicate issue as remarked earlier. In addition to the already mentioned works [3, 12, 15], see also [6, 9], and references therein.

We are now ready to show the following. We take $\boldsymbol{u}_\circ$ as a bi-Lipschitz mapping between two Lipschitz, bounded domains

$$\Omega, \quad \Omega_\circ = \boldsymbol{u}_\circ(\Omega),$$

in such a way that

$$\boldsymbol{u}_\circ \in W^{1,\infty}(\Omega; \Omega_\circ), \quad \boldsymbol{u}_\circ^{-1} \in W^{1,\infty}(\Omega_\circ; \Omega).$$

Theorem 3.2 *Suppose that the density for hyper-elasticity $W(\mathbf{X}) : \mathbb{R}_+^{3\times3} \to \mathbb{R}$ is poly-convex, and*

$$W(\mathbf{X}) \geq C \left(\frac{1}{\det \mathbf{X}} |\mathbf{X}|^2 + \frac{1}{\det \mathbf{X}} |\operatorname{adj} \mathbf{X}|^2 + \frac{1}{\det \mathbf{X}^{r-1}} + |\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r \right),$$

for some positive constant $C > 0$, exponent $r > 1$, and that boundary data are provided by a mapping $\boldsymbol{u}_\circ$ as just indicated. The two variational problems (2.3)-(2.4), and (2.5)-(2.6) admit minimizers, and are equivalent in the sense that rule (3.1) permits to pass from minimizers of one to minimizers of the other.

Proof By Theorems 2.1 and 2.5, problems (2.3) and (2.5) admit minimizers, given that hypotheses in the statement guarantee assumptions in those respective results.

(1) Suppose $\boldsymbol{u} \in \mathcal{A}_+$ for E in (2.3). Since

$$W(\mathbf{X}) \geq C \frac{1}{\det \mathbf{X}} |\operatorname{adj} \mathbf{X}|^2 = C \det \mathbf{X} |\mathbf{X}^{-1}|^2,$$

we can conclude that

$$\int_\Omega \det \nabla \boldsymbol{u}(\boldsymbol{x}) |\nabla \boldsymbol{u}(\boldsymbol{x})^{-1}|^2 \, d\boldsymbol{x} \leq \frac{1}{C} \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x})) \, d\boldsymbol{x} < \infty.$$

By Lemma 3.1, we see that $\boldsymbol{u}^{-1} \in H^1(\boldsymbol{u}(\Omega); \mathbb{R}^3)$, and, then, the mapping $\Phi = \boldsymbol{u}^{-1} \circ \boldsymbol{u}_\circ$ given in (3.1) belongs to $\mathcal{I}_+$ with $I(\Phi) = E(\boldsymbol{u})$.

(2) Conversely, let $\Phi \in \mathcal{I}_+$ and I given in (2.5). As before, put $\mathbb{H}$ for the compact set of matrices in $\mathbb{R}_+^{3\times3}$ given by

$$\mathbb{H} = \{\nabla \boldsymbol{u}_\circ(\boldsymbol{x}) : \boldsymbol{x} \in \Omega\}.$$

Since all matrices involved, $\nabla u_\circ(x)$ and $\nabla \Phi(x)$, have positive determinant, there is a constant $\tilde{C} > 0$ (depending on C and $\mathbb{H}$) such that

$$W(\mathbf{FX}) \geq C|\mathbf{FX}|^2 \geq \tilde{C}|\mathbf{X}|^2.$$

Note that again our lower bound on W ensures trivially that

$$W(\mathbf{Y}) \geq C|\mathbf{Y}|^2.$$

Hence, the integral

$$\int_\Omega |\nabla\Phi(x)^{-1}|^2 \det \nabla\Phi(x)\,dx \leq \frac{1}{C}\int_\Omega W(\nabla u_\circ(x)\nabla\Phi(x)^{-1}) \det \nabla\Phi(x)\,dx$$

is finite, and by Lemma 3.1, $\Phi^{-1} \in H^1(\Omega; \mathbb{R}^3)$. Again (3.1) ensures that

$$u = u_\circ \circ \Phi^{-1} \in \mathcal{A}_+ \text{ and } E(u) = I(\Phi).$$

Because of the arbitrariness of $u \in \mathcal{A}_+$ and $\Phi \in \mathcal{I}_+$, this discussion finishes the proof. $\qquad\square$

From now on, we will take for granted that densities $W(\mathbf{X})$ for hyperelasticity comply with the two fundamental hypotheses in this theorem.

4 Optimality

The derivation of optimality conditions in hyper-elasticity is exposed to important difficulties as has been stressed in the Introduction, and it is well-known in the field. The use of the standard form of making additive variations

$$u(x) + \epsilon U(x), \quad U(x) = 0 \text{ on } \partial\Omega, \tag{4.1}$$

stumbles with the fundamental issue of ensuring that the maps in this family are one-to-one for small ϵ. It is not even clear how to enforce

$$\det \nabla(u + \epsilon U) > 0$$

without asking for the uniform positivity of the jacobian for the minimizer u. Said differently, variations of the form (4.1) can hardly be shown to belong to either $\mathcal{A}_+$ or $\mathcal{I}_+$ even for small ϵ.

We thus face the issue of producing one-parameter, smooth families $u_\epsilon(x)$ of maps belonging to either $\mathcal{A}_+$ or $\mathcal{I}_+$ for every small (in absolute value) ϵ with u_0 (u_ϵ for $\epsilon = 0$) being the minimizer of E, or representing, through (3.1), the minimizer for I in $\mathcal{I}_+$, respectively. To this end, we concentrate on $\mathcal{I}_+$ as this is one of the main motivations to explore the inner version.

Our proposal is to use the flow associated with test infinitesimal generators Ψ to produce feasible, continuous families of maps belonging to $\mathcal{I}_+$. This was already indicated in [18]. Namely, take a smooth

$$\Psi(z) : \Omega \to \mathbb{R}^3, \quad \Psi(z) = 0 \text{ for } z \in \partial\Omega, \tag{4.2}$$

and let

$$\Phi(t; \boldsymbol{x}) : \mathbb{R} \times \Omega \to \Omega$$

be the flow associated with Ψ, i.e.

$$\Phi'(t; \boldsymbol{x}) = \Psi(\Phi(t; \boldsymbol{x})), \quad (t, \boldsymbol{x}) \in \mathbb{R} \times \Omega, \quad \Phi(0; \boldsymbol{x}) = \boldsymbol{x}. \tag{4.3}$$

Proposition 4.1 *For every smooth Ψ as in (4.2), its flow $\Phi(t; \boldsymbol{x})$ is smooth (in all its variables) and such that $\Phi(t; \cdot) \in \mathcal{I}_+$ (recall (2.2)). If $\boldsymbol{u} \in \mathcal{A}_+$, then $\boldsymbol{u}(\Phi(t; \cdot)) \in \mathcal{A}_+$ for all real $t \in \mathbb{R}$.*

Proof The proof is standard. Each individual

$$\Phi(t; \cdot) : \Omega \to \Omega$$

is a feasible map for our variational problem because $\Phi(-t; \cdot)$ is its inverse due to the semigroup property; $\Phi(t; \boldsymbol{x}) = \boldsymbol{x}$ for all $t \in \mathbb{R}$ and $\boldsymbol{x} \in \partial\Omega$; and the dependence of $\Phi(t; \cdot)$ on $\boldsymbol{x}$ is smooth due to the dependence of solutions of ODEs with respect to initial conditions and parameters. $\square$

By Theorem 3.2, minimizers for I and for E are related through (3.1), and so we can use feasible variations in Proposition 4.1 to derive optimality conditions for both E and I. The manipulations that follow are inspired by the similar ones in [5] which were performed in the context of standard variations (4.1).

Suppose $\boldsymbol{u}$ is a minimizer for E in $\mathcal{A}_+$. By Proposition 4.1, we should formally have

$$\frac{d}{dt}\bigg|_{t=0} \int_\Omega W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x} = 0, \tag{4.4}$$

given that $\Phi(0; \boldsymbol{x})) = \boldsymbol{x}$ in Ω. Unless we assume either that the minimizer $\boldsymbol{u}$ is Lipschitz or the uniform positivity of the determinant $\det \nabla \boldsymbol{u}$ in Ω, the passage of the previous differentiation under the integral sign is hard to justify. Yet we claim that, under some convenient additional assumptions, we should indeed have

$$\int_\Omega \frac{d}{dt}\bigg|_{t=0} W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x} = 0, \tag{4.5}$$

even if the function

$$t \mapsto g(t) \equiv \int_\Omega W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x}$$

does not admit derivative at $t = 0$.

Lemma 4.2 *Suppose $\boldsymbol{u}$ is a minimizer for E in $\mathcal{A}_+$. If for smooth Ψ as in (4.2), and corresponding flow $\Phi(t; \boldsymbol{x})$, the function*

$$t \in \mathbb{R} \mapsto G(t; \boldsymbol{x}) \equiv W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x})))$$

is differentiable for a.e. $\boldsymbol{x} \in \Omega$, and there is a $L^1(\Omega)$-function $\overline{W}(\boldsymbol{x})$ such that

$$\left|\frac{d}{dt}G(t; \boldsymbol{x})\right| \leq \overline{W}(\boldsymbol{x}) \tag{4.6}$$

in a vicinity of $t = 0$, then

$$\int_\Omega \frac{d}{dt}\Big|_{t=0} W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x} = 0.$$

Proof For a big, finite $L > 0$, put

$$g_L(t) = \int_{\Omega_L} W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x}, \quad \Omega_L = \{\boldsymbol{x} \in \Omega : |\nabla\boldsymbol{u}(\boldsymbol{x})| \leq L\}.$$

We claim that, because of (4.6):

(1) for each fixed L, $g_L(t)$ is differentiable because $\boldsymbol{u}$ is Lipschitz in Ω_L and by dominated convergence

$$g_L'(t) = \int_{\Omega_L} \frac{d}{dt} W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x};$$

(2) $g_L(t) \to g(t)$ as $L \to \infty$ for every t, by monotone convergence given that $W \geq 0$.

Select $t_+ > 0$, $t_- < 0$ in the vicinity of $t = 0$ assumed in the statement, such that

$$g(t_+) > g(0), \quad g(t_-) > g(0).$$

This is possible because g has a global minimum at $t = 0$, and g is non-constant in a neighborhood of $t = 0$ (otherwise there would be nothing to be shown). For L sufficiently large, by the convergence written above $g_L(t) \to g(t)$,

$$g_L(t_+) > g_L(0), \quad g_L(t_-) > g_L(0).$$

By Rolle's theorem, since g_L is differentiable, there is $t_L \in (t_-, t_+)$ such that $g_L'(t_L) = 0$, i.e.

$$\int_{\Omega_L} \frac{d}{dt}\Big|_{t=t_L} W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x} = 0.$$

By the upper bound assumed in the statement and dominated convergence, we can pass to the limit in L, and conclude that

$$\int_\Omega \frac{d}{dt}\Big|_{t=t^*} W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x} = 0,$$

for some $t^* \in (t_-, t_+)$. Since the interval (t_-, t_+) can be made arbitrarily small around $t = 0$, the proof is finished. $\qquad\square$

The technical condition in this statement essentially guarantees, after all, that even if the function g is not differentiable, the function resulting in taking the t-derivative under the integral sign is integrable so that the optimality condition is legitimate.

In computing the derivative of the auxiliary function $G(t; \boldsymbol{x})$ with respect to t in (4.5), we need to rely on the differentiability of the integrand W to utilize the chain rule in a usual way. Formally, we would find

$$G'(t; \boldsymbol{x}) = \nabla W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \nabla (\nabla\boldsymbol{u}(\Phi(-t; \boldsymbol{x})) \Psi(\Phi(-t; \boldsymbol{x}))),$$

where we have taken into account (4.3); and (4.5) becomes, by letting $t = 0$ and bearing in mind initial conditions in (4.3),

$$\int_\Omega \nabla W(\nabla \boldsymbol{u}(\boldsymbol{x})) \nabla [\nabla \boldsymbol{u}(\boldsymbol{x}) \Psi(\boldsymbol{x})] \, d\boldsymbol{x} = 0 \tag{4.7}$$

for all such test fields Ψ. This optimality condition on the minimizer $\boldsymbol{u}$ is the weak form of the classic Noether conservation law corresponding to the appropriate invariance (see for instance [11]). Under hypotheses permitting to interpret

$$\Theta(\boldsymbol{x}) = \nabla \boldsymbol{u}(\boldsymbol{x}) \Psi(\boldsymbol{x})$$

as a legitimate test field, we would have

$$\int_\Omega \nabla W(\nabla \boldsymbol{u}(\boldsymbol{x})) \nabla \Theta(\boldsymbol{x}) \, d\boldsymbol{x} = 0.$$

Even further, if the class of test fields Θ of the previous form can cover all possible test fields, or at least a dense subset of them, then we would recover the usual weak form of the Euler-Lagrange system of optimality. That property would however require to put on the minimizer $\boldsymbol{u}$ sufficient regularity for the inverse mapping $\boldsymbol{u}^{-1}$ to be defined and enjoy regularity properties beyond the ones assumed in Theorem 3.2. Check [18] for one particular result in this direction which is also stated in the framework (4.2)-(4.3).

Another possibility is to exploit again the passage from E to I in (4.4) before taking any derivative. As before, we find that

$$\int_\Omega W(\nabla(\boldsymbol{u} \circ \Phi(-t; \boldsymbol{x}))) \, d\boldsymbol{x} = \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x}) \nabla \Phi(t; \boldsymbol{x})^{-1}) \det \nabla \Phi(t; \boldsymbol{x}) \, d\boldsymbol{x}.$$

To carry the derivative of the resulting function

$$t \mapsto \int_\Omega W(\nabla \boldsymbol{u}(\boldsymbol{x}) \nabla \Phi(t; \boldsymbol{x})^{-1}) \det \nabla \Phi(t; \boldsymbol{x}) \, d\boldsymbol{x}$$

into the integral sign looks easier to justify than in the previous case as the variable t does not interfere with the minimizer $\boldsymbol{u}$. However, as a matter of fact, one needs to use a similar argument as in the proof of Lemma 4.2: it should be required that the outcome of the differentiation under the integral sign be integrable, namely the function $\mathbf{G}(t; \cdot)$ given by

$$W(\nabla \boldsymbol{u}(\cdot) \nabla \Phi(t; \cdot)^{-1}) \operatorname{adj} \nabla \Phi(t; \cdot) \nabla \Psi(\Phi(t; \cdot)) - \tag{4.8}$$

$$\nabla \Phi(t; \cdot)^{-T} \nabla \boldsymbol{u}(\cdot)^T \nabla W(\nabla \boldsymbol{u}(\cdot) \nabla \Phi(t; \cdot)^{-1}) \nabla \Psi(\Phi(t; \cdot)) \operatorname{adj} \nabla \Phi(t; \cdot),$$

ought to be uniformly integrable in a neighborhood of $t = 0$. We provide next a set of conditions ensuring this assumption. The tensor involved in this form of optimality is the energy-momentum tensor, the multidimensional counterpart of the Du Bois Reymond of Erdmann equations.

Theorem 4.3 *Suppose that the integrand W for (2.3) is differentiable (whenever finite), and for every compact set $\mathbb{K} \subset \mathbb{R}_+^{3 \times 3}$ of matrices with positive determinant, there is a positive constant $C \equiv C(\mathbb{K})$ such that*

$$|\mathbf{B}^T \mathbf{A}^T \nabla W(\mathbf{A}\mathbf{B})| \leq C(1 + W(\mathbf{A})), \quad |W(\mathbf{A}\mathbf{B})| \leq C(1 + W(\mathbf{A})),$$

for $\mathbf{A} \in \mathbb{R}_+^{3\times3}$ and $\mathbf{B} \in \mathbb{K}$. Let $\boldsymbol{u} \in \mathcal{A}_+$ be a minimizer for (2.3) in $\mathcal{A}_+$. Then, for every smooth Ψ as in (4.2), we should have

$$\int_\Omega [W(\nabla \boldsymbol{u}(\boldsymbol{x}))\mathbf{id} - \nabla \boldsymbol{u}(\boldsymbol{x})^T \nabla W(\nabla \boldsymbol{u}(\boldsymbol{x}))]\nabla \Psi(\boldsymbol{x}) \, d\boldsymbol{x} = 0.$$

Proof The proof has almost been indicated above. The new upper bounds assumed on W for the compact set

$$\mathbb{K} = \{\nabla \Phi(t; \boldsymbol{x})^{-1} : \boldsymbol{x} \in \Omega, |t| \leq \epsilon_0\}, \quad \epsilon_0 > 0,$$

imply that the function $\mathbf{G}(t; \cdot)$ in (4.8) is bounded above by

$$\tilde{C}(1 + W(\nabla \boldsymbol{u}(\boldsymbol{x})))$$

which is a fixed $L^1(\Omega)$-function. The constant $\tilde{C}$ here involves also a uniform constant for the products

$$|\operatorname{adj} \nabla \Phi(t; \boldsymbol{x})| \, |\nabla \Psi(\Phi(t; \boldsymbol{x}))|, \quad \boldsymbol{x} \in \Omega, |t| \leq \epsilon_0.$$

Note that the fields $\Psi(\boldsymbol{x})$ and $\Phi(t; \boldsymbol{x})$ are smooth, and hence uniformly bounded, in $\boldsymbol{x}$ and $(t, \boldsymbol{x})$, respectively. $\qquad\square$

The upper bounds in this statement do not typically occur in existence theorems in hyper-elasticity. They rather occur in relaxation facts ([8]), or in connection with optimality ([5]).

Remark 4.4 Note that condition (4.2) can be relaxed to

$$\Psi(\boldsymbol{z}) : \Omega \to \mathbb{R}^3, \quad \Psi(\boldsymbol{z}) \cdot \mathbf{n}(\boldsymbol{x}) = 0,$$

either in all or part of the boundary $\partial\Omega$, to deal with pure traction problems or mixed cases. $\mathbf{n}$ is the outer, unit normal to $\partial\Omega$.

5 The Outer Version

We focus now in problem (1.9) for a feasible set given in (1.12), namely

$$\text{Minimize in } \Psi \in \mathcal{O}_+: \quad O(\Psi) = \int_\Omega W(\nabla(\Psi \circ \boldsymbol{u}_\circ)(\boldsymbol{x})) \, d\boldsymbol{x}$$

where $\mathcal{O}_+$ is defined by

$$\mathcal{O}_+ = \{\Psi \in H^1(\Omega_\circ; \mathbb{R}^3) : \Psi = \mathbf{id} \text{ on } \Gamma_\circ, \Psi \text{ is one-to-one a.e. in } \Omega_\circ, \det \nabla \Psi > 0 \text{ a.e. in } \Omega_\circ\}.$$

Recall that $\Omega_\circ = \boldsymbol{u}_\circ(\Omega)$.

In this complementary setting, the roles of feasible Ψ's and fixed $\boldsymbol{u}_\circ$ are interchanged with respect to the inner version. Because of this reason the outer version is much less informative. We can write, just as before,

$$O(\Psi) = \int_{\Omega_\circ} W(\nabla \Psi(\boldsymbol{x}) \nabla \boldsymbol{u}_\circ^{-1}(\boldsymbol{x})^{-1}) \det \nabla \boldsymbol{u}_\circ^{-1}(\boldsymbol{x}) \, d\boldsymbol{x}. \qquad (5.1)$$

This formulation, formally equivalent to the initial one, does not provide more information on W because $\boldsymbol{u}_{\circ}$ is a fixed deformation. Check [16] for a similar situation in which a quite involved optimal control problem in hyper-elasticity in the context of differential growth is examined. Indeed, one can show existence of global minimizers for this outer version under the usual assumptions on W. The following result is similar to Theorem 2.1. It is just the typical framework in hyper-elasticity.

Proposition 5.1 *Suppose that the function $W(\mathbf{X})$ is a poly-convex density for hyper-elasticity with*

$$W(\mathbf{X}) \geq C(|\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r), \quad C > 0, r > 1, \mathbf{X} \in \mathbb{R}^{3 \times 3}_+.$$

If $\boldsymbol{u}_{\circ}$ is a bi-Lipschitz mapping from Ω to $\Omega_{\circ} = \boldsymbol{u}_{\circ}(\Omega)$, then there are minimizers Ψ for functional $O(\Psi)$ in (5.1) over $\mathcal{O}_+$.

Proof The proof is standard. Let $\mathbf{F}$ stand for a generic matrix in the compact set

$$\mathbb{K} = \{\nabla \boldsymbol{u}_{\circ}^{-1}(\boldsymbol{x})^{-1} : \boldsymbol{x} \in \Omega_{\circ}\}.$$

After the proof of Lemma 2.3, it is now elementary to check that the integrand W is poly-convex if and only if the family of integrands in the variable $\mathbf{X}$ defined by

$$W_{\mathbf{F}}(\mathbf{X}) = \frac{1}{\det \mathbf{F}} W(\mathbf{XF})$$

for fixed matrix $\mathbf{F} \in \mathbb{K}$, are poly-convex. On the other hand, for every such $\mathbf{F} \in \mathbb{K}$,

$$W_{\mathbf{F}}(\mathbf{X}) \geq \frac{C}{\det \mathbf{F}}(|\mathbf{XF}|^2 + |\operatorname{adj}(\mathbf{XF})|^2 + \det(\mathbf{XF})^r).$$

As we used earlier,

$$|\mathbf{XF}| \geq |\mathbf{F}^{-1}|^{-1}|\mathbf{X}|, \quad |\operatorname{adj}(\mathbf{XF})| \geq |\operatorname{adj} \mathbf{F}^{-1}|^{-1}|\operatorname{adj} \mathbf{X}|.$$

Hence

$$W_{\mathbf{F}}(\mathbf{X}) \geq \frac{C}{\det \mathbf{F}}(|\mathbf{F}^{-1}|^{-2}|\mathbf{X}|^2 + |\operatorname{adj} \mathbf{F}^{-1}|^{-2}|\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{F}^r \det \mathbf{X}^r).$$

There is some positive constant $\tilde{C}$ (depending on C and $\mathbb{K}$) such that

$$W_{\mathbf{F}}(\mathbf{X}) \geq \tilde{C}(|\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r)$$

uniformly for $\mathbf{F} \in \mathbb{K}$. The existence of minimizers in $\mathcal{O}_+$ is a direct consequence of Theorem 2.1 □

Optimality conditions can be analyzed in this outer framework much in the same way as we did in the previous section. The conditions that can be derived are even more involved (see [5]) though they can be rigorously deduced within the context of one-parameter families belonging truly to $\mathcal{O}_+$ as in the last section.

6 Augmented Hyper-Elastic Materials

In this final section, we would like to explore how the lower bound

$$W(\mathbf{X}) \geq C \left(\frac{1}{\det \mathbf{X}} |\mathbf{X}|^2 + \frac{1}{\det \mathbf{X}} |\operatorname{adj} \mathbf{X}|^2 + \frac{1}{\det \mathbf{X}^{r-1}} + |\mathbf{X}|^2 + |\operatorname{adj} \mathbf{X}|^2 + \det \mathbf{X}^r \right),$$

in Theorem 3.2 can be the basis of explicit valid densities for some hyper-elastic materials. In general, given a valid internal energy density $W(\mathbf{X}) : \mathbb{R}^{3\times3}_+ \to \mathbb{R}$ to which Theorem 2.1 can be applied, Theorem 3.2 can be utilized for the "augmented" energy density

$$AW(\mathbf{X}) = W(\mathbf{X}) + \mathbb{T}W(\mathbf{X}) = W(\mathbf{X}) + \det \mathbf{X} W(\mathbf{X}^{-1}), \tag{6.1}$$

where $\mathbb{T}$ is given in (1.16), provided boundary data are suitably given. In particular, there are minimizers for the internal energy functional

$$\int_{\Omega} AW(\nabla \boldsymbol{u}(\boldsymbol{x})) \, d\boldsymbol{x}$$

in the appropriate class, complying with suitable boundary data. Equivalently, if further regularity is known or imposed on mapping $\boldsymbol{u}_\circ$, then such minimizers can be understood too in the form of changes of variables (inner mappings).

Frame indifference and isotropy are easily handled.

Lemma 6.1 *Suppose W is a valid energy density for hyper-elasticity, and let $AW(\mathbf{X})$ be given in (6.1).*

(1) *If W is frame-indifferent, and $W(\mathbf{X}) = W(\mathbf{X}^T)$ for all $\mathbf{X} \in \mathbb{R}^{3\times3}_+$, then AW is frame-indifferent.*
(2) *If W is isotropic, and $W(\mathbf{X}) = W(\mathbf{X}^T)$ for all $\mathbf{X} \in \mathbb{R}^{3\times3}_+$, then AW is isotropic.*
(3) *If W is frame-indifferent, and isotropic, then AW is frame-indifferent, and isotropic.*

Proof The proofs of these statements are elementary. For the first, simply note that for a rotation $\mathbf{Q}$, $\mathbf{Q}^T\mathbf{Q} = \mathbf{id}$, $\det \mathbf{Q} = 1$,

$$\begin{aligned} AW(\mathbf{QX}) &= W(\mathbf{QX}) + \det \mathbf{X} W(\mathbf{X}^{-1}\mathbf{Q}^T) \\ &= W(\mathbf{X}) + \det \mathbf{X} W(\mathbf{QX}^{-T}) \\ &= W(\mathbf{X}) + \det \mathbf{X} W(\mathbf{X}^{-T}) \\ &= W(\mathbf{X}) + \det \mathbf{X} W(\mathbf{X}^{-1}) \\ &= AW(\mathbf{X}). \end{aligned}$$

We have used several times the frame-indifference of W and its symmetric invariance. The second statement is similar. For the third one, go back to the first line of the previous chain of equalities

$$AW(\mathbf{QX}) = W(\mathbf{QX}) + \det \mathbf{X} W(\mathbf{X}^{-1}\mathbf{Q}^T)$$

and use frame-indifference and isotropy for W, respectively, in the two terms on the right-hand side, to conclude that

$$AW(\mathbf{QX}) = W(\mathbf{X}) + \det \mathbf{X} W(\mathbf{X}^{-1}).$$

Argue in a similar way to show isotropy for AW. $\qquad\square$

The behavior as $\det \mathbf{X} \to 0^+$ is also elementary.

Lemma 6.2 *Suppose $W(\mathbf{X}) : \mathbb{R}^{3\times 3} \to \mathbb{R}$ is non-negative and such that*

$$\lim_{\det \mathbf{X} \to \infty} \frac{W(\mathbf{X})}{\det \mathbf{X}} = \infty.$$

Then $AW(X)$ is a valid density for hyper-elasticity, i.e.

$$\lim_{\det \mathbf{X} \to 0^+} AW(\mathbf{X}) = \infty.$$

Proof The proof is straightforward under the change of variables

$$\mathbf{X} \mapsto \mathbf{Y} = \mathbf{X}^{-1}$$

in the first limit. □

As an explicit example of this process we consider the family of stored energy densities

$$W_\circ(\mathbf{X}) = \frac{a}{\det \mathbf{X}}|\mathbf{X}|^2 + \frac{b}{\det \mathbf{X}}|\operatorname{adj}\mathbf{X}|^2 + c\frac{1}{\det \mathbf{X}^{r-1}} + d|\mathbf{X}|^2 + e|\operatorname{adj}\mathbf{X}|^2 + f\det \mathbf{X}^r, \quad (6.2)$$

defined for $\det \mathbf{X} > 0$, for positive constants a, b, c, d, e, f, and exponent $r > 1$. The following is a corollary of our previous discussions. Note that

$$W_\circ(\mathbf{X}) \to +\infty \text{ as } \det \mathbf{X} \to 0^+.$$

This is in fact a consequence of Lemma 6.2. Ogden materials ([7, 17]) provide a more complete family of energy densities to which this process can be applied. Some similar examples of poly-convex functions were considered in [1].

Theorem 6.3 *The energy density $W_\circ$ just defined is valid for hyper-elasticity, isotropic, and poly-convex. The variational problem*

$$\text{Minimize in } \boldsymbol{u} \in \mathcal{A}_+ : \quad E(\boldsymbol{u}) = \int_\Omega W_\circ(\nabla \boldsymbol{u}(\boldsymbol{x}))\,d\boldsymbol{x}$$

for

$$\mathcal{A}_+ = \{\boldsymbol{u} \in H^1(\Omega; \mathbb{R}^3) : \boldsymbol{u} = \boldsymbol{u}_\circ \text{ on } \Gamma, \boldsymbol{u} \text{ is one-to-one a.e. in } \Omega, \det \nabla \boldsymbol{u} > 0 \text{ a.e. in } \Omega\}$$

for suitable $\boldsymbol{u}_\circ$, admits global minimizers.

A more general family of energy densities for hyper-elasticity is that of Mooney-Rivlin materials

$$W(\mathbf{X}) = a|\mathbf{X}|^2 + b|\operatorname{adj}\mathbf{X}|^2 + g(\det \mathbf{X}), \quad a, b > 0. \tag{6.3}$$

Contrary to the function $g(t) = t^r$, with $r > 1$, the function g is usually assumed to have an infinite limit as $\det \mathbf{X} \to 0^+$. One possibility here is to ask for the condition

$$\lim_{t \to \infty} \frac{g(t)}{t} = \infty,$$

and then replace $g(\det\mathbf{X})$ in (6.3) by

$$g(\det\mathbf{X}) + \det\mathbf{X}\, g(\det\mathbf{X}^{-1})$$

to ensure the appropriate behavior as $\det\mathbf{X} \to 0^+$. Often, the function $g(t)$ is chosen in the form

$$g(t) = \frac{\lambda}{2}(t-1)^2 - \mu\log t, \quad \lambda, \mu > 0.$$

In this case the term $tg(1/t)$ becomes

$$\frac{\lambda}{2}\frac{(t-1)^2}{t} + \mu t\log t,$$

and the sum of the two terms becomes

$$\frac{\lambda}{2}(t-1)^2\frac{t+1}{t} + \mu(t-1)\log t,$$

which has a vanishing global minimum at $t=1$ for every positive λ and μ.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

1. Bandeira, L., Pedregal, P.: $\mathcal{A}$-variational principles. Milan J. Math. **91**, 293–314 (2023)
2. Ball, J.M.: Convexity conditions and existence theorems in nonlinear elasticity. Arch. Ration. Mech. Anal. **63**(4), 337–403 (1976)
3. Ball, J.M.: Global invertibility of Sobolev functions and the interpenetration of matter. Proc. R. Soc. Edinb., Sect. A **88**(3–4), 315–328 (1981)
4. Ball, J.M.: Minimizers and the Euler-Lagrange equations. In: Lecture Notes in Phys., vol. 195, pp. 1–4. Springer, Berlin (1984)
5. Ball, J.M.: Some open problems in elasticity. In: Geometry, Mechanics, and Dynamics, pp. 3–59. Springer, New York (2002)
6. Barchiesi, M., Henao, D., Mora-Corral, C.: Local invertibility in Sobolev spaces with applications to nematic elastomers and magnetoelasticity. Arch. Ration. Mech. Anal. **224**(2), 743–816 (2017)

7. Ciarlet, P.G.: Mathematical Elasticity. Vol. I. Three-Dimensional Elasticity. Studies in Mathematics and Its Applications, vol. 20. North-Holland, Amsterdam (1988)
8. Conti, S., Dolzmann, G.: On the theory of relaxation in nonlinear elasticity with constraints on the determinant. Arch. Ration. Mech. Anal. **217**(2), 413–437 (2015)
9. Fonseca, I., Gangbo, W.: Local invertibility of Sobolev functions. SIAM J. Math. Anal. **26**(2), 280–304 (1995)
10. Dacorogna, B.: Direct Methods in the Calculus of Variations, 2nd edn. Applied Mathematical Sciences, vol. 78. Springer, New York (2008)
11. Giaquinta, M., Hildebrandt, S.: Calculus of Variations. I. The Lagrangian Formalism. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 310. Springer, Berlin (1996)
12. Henao, D., Mora-Corral, C.: Invertibility and weak continuity of the determinant for the modelling of cavitation and fracture in nonlinear elasticity. Arch. Ration. Mech. Anal. **197**(2), 619–655 (2010)
13. Iwaniec, T., Onninen, J.: The Dirichlet principle for inner variations. Math. Ann. **383**(1–2), 315–351 (2022)
14. Iwaniec, T., Martin, G., Onninen, J.: Energy-minimal principles in geometric function theory. N.Z. J. Math. **52**, 605–642 (2021)
15. Müller, S., Qi, T., Yan, B.S.: On a new class of elastic deformations not allowing for cavitation. Ann. Inst. Henri Poincaré, Anal. Non Linéaire **11**(2), 217–243 (1994)
16. Ortigosa-Martínez, R., Martínez-Frutos, J., Mora-Corral, C., Pedregal, P., Periago, F.: Shape-programming in hyper-elasticity through differential growth. Appl. Math. Optim. https://doi.org/10.1007/s00245-024-10117-6
17. Pedregal, P.: Variational Methods in Nonlinear Elasticity. SIAM, Philadelphia (2000)
18. Pedregal, P.: A multiplicative version of quasiconvexity for hyperelasticity. J. Elast. **151**(2), 219–236 (2022)
19. Pourciau, B.H.: Univalence and degree for Lipschitz continuous maps. Arch. Ration. Mech. Anal. **81**(3), 289–299 (1983)